Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **498,** 1852–1870 (2020) doi:10.1093/mnras/staa2504
Advance Access publication 2020 August 20

Do model emission line galaxies live in filaments at $z \sim 1$?

V. Gonzalez-Perez [1,2,3★] W. Cui [4,5] S. Contreras [6] C. M. Baugh,[7] J. Comparat,[8] A. J. Griffin [7] J. Helly,[7] A. Knebe [4,9,10] C. Lacey [7] and P. Norberg[7,11]

Affiliations are listed at the end of the paper

Accepted 2020 August 12. Received 2020 August 11; in original form 2020 January 17

ABSTRACT

Current and future cosmological surveys are targeting star-forming galaxies at $z \sim 1$ with nebular emission lines. We use a state-of-the-art semi-analytical model of galaxy formation and evolution to explore the large-scale environment of star-forming emission line galaxies (ELGs). Model ELGs are selected such that they can be compared directly with the DEEP2, VVDS, eBOSS-SGC, and DESI surveys. The large-scale environment of the ELGs is classified using velocity–shear–tensor and tidal–tensor algorithms. Half of the model ELGs live in filaments and about a third in sheets. Model ELGs that reside in knots have the largest satellite fractions. We find that the shape of the mean halo occupation distribution of model ELGs varies widely for different large-scale environments. To interpret our results, we also study fixed number density samples of ELGs and galaxies selected using simpler criteria, with single cuts in stellar mass, star formation rate, and [O II] luminosity. The fixed number density ELG selection produces samples that are close to L[O II] and SFR-selected samples for densities above $10^{-4.2}\,h^3\,\mathrm{Mpc}^{-3}$. ELGs with an extra cut in stellar mass applied to fix their number density, present differences in sheets and knots with respect to the other samples. ELGs, SFR, and L[O II] selected samples with equal number density have similar large-scale bias but their clustering below separations of $1h^{-1}$ Mpc is different.

Key words: galaxies: evolution – galaxies: formation – large-scale structure of Universe.

1 INTRODUCTION

The distribution of matter in the Universe is highly inhomogeneous on megaparsec scales, on which the filamentary structure of the cosmic web arises (e.g. Colless et al. 2001; Gott et al. 2005; Cui et al. 2018). In the prevalent theory of hierarchical formation of structure, gas cools following the cosmic web (e.g. White & Rees 1978). Galaxies at different distances from the filamentary structures have been found to have different properties that cannot be explained by density alone (e.g. Kraljic et al. 2018; Laigle et al. 2018), although this might not be a universal result (e.g. Goh et al. 2019).

Here, we aim to study the large-scale environment of star-forming emission line galaxies (hereafter ELGs). These galaxies have spectra characterized by strong nebular emission lines, which allow the robust determination of their redshift. Cosmological surveys have started to target ELGs to study the epoch when the expansion of the Universe first became dominated by dark energy, $z \sim 1$ (e.g. Comparat et al. 2013a). Understanding the connection between ELGs and their host dark matter haloes is a crucial step to maximally exploit these surveys. Cosmological surveys that have targetted or plan to target ELGs include ATLAS Probe,[1] DESI,[2] *Euclid*,[3] Hetdex,[4] MSE,[5] PFS,[6] SDSS-IV/eBOSS,[7] WFIRST[8] WiggleZ,[9] 4MOST,[10] etc. Most of these surveys have either optical or infrared detectors and thus, they will focus on detecting H$\alpha\lambda$6563 Å, H$\beta\lambda$4861 Å, [O II] $\lambda\lambda$3726, 3729 Å, and [O III] $\lambda\lambda$4959, 5007 Å nebular emission lines at redshifts between 0.5 and 2 (e.g. Sobral et al. 2012). Here, we focus on [O II] emitters, which are the prevalent ELGs detected at $z \sim 1$ by optical instruments, such as those from SDSS-IV/eBOSS (Comparat et al. 2016b; Delubac et al. 2017; Raichoor et al. 2017).

The nebular emission lines in ELGs are produced by ionized gas in the interstellar medium. The gas can be heated by either newly formed stars or by the nuclear activity following mass accretion into an SMBH. In this study, we focus only on star-forming ELGs, as this is the population targeted by the cosmological surveys mentioned above. Only a small fraction, less than 1 per cent, of eBOSS ELGs are expected to be active galactic nuclei (AGNs; Comparat et al. 2013b). Other observational studies have found that between 8 per cent and

★ E-mail: violetagp@protonmail.com

[1]Astrophysics Telescope for Large Area Spectroscopy, http://atlas-probe.ipac.caltech.edu/ (Wang et al. 2018).
[2]Dark Energy Spectroscopic Instrument, http://desi.lbl.gov/ (Levi et al. 2013).
[3]https://www.euclid-ec.org/ (Laureijs et al. 2011).
[4]Hobby-Eberly Telescope Dark Energy Experiment, http://hetdex.org/ (Hill et al. 2008).
[5]MaunaKea Spectroscopic Explorer, https://mse.cfht.hawaii.edu/ (Percival et al. 2019).
[6]Prime Focus Spectrograph, http://sumire.ipmu.jp/en/2652 (Takada et al. 2014).
[7]Extended Baryon Oscillation Spectroscopic Survey, http://www.sdss.org/surveys/eboss/ (Ahumada et al. 2019).
[8]Wide Field Infrared Survey Telescope, https://www.nasa.gov/wfirst (Hounsell et al. 2018).
[9]WiggleZ Dark Energy Survey, http://wigglez.swin.edu.au/site/ (Drinkwater et al. 2010, 2018).
[10]4-metre Multi-Object Spectroscopic Telescope, https://www.4most.eu/ (de Jong et al. 2014).

17 per cent of their ELGs were AGNs (Sobral et al. 2016; Valentino et al. 2017). We will refer to star-forming ELGs simply as ELGs and thus, here we study a subsample of star-forming galaxies.

Star-forming and less massive galaxies have been found closer to axis of filaments than more massive or quiescent galaxies are by a range of observational studies including GAMA at $0.01 \leq z \leq 0.25$ (Alpaslan et al. 2016; Kraljic et al. 2018), CHILES at $z \sim 0.45$ (Luber et al. 2019) SDSS at $z \leq 0.7$ (Chen et al. 2017, Poudel et al. 2017), VIPERS at $z \sim 0.7$ (Malavasi et al. 2017), VIS^{3}COS at $z \sim 0.89$ (Paulino-Afonso et al. 2019), COSMOS at $z < 0.9$ (Laigle et al. 2018), etc. The HiZELS survey also found that galaxies in filaments at $z \sim 0.53$ and $z \sim 0.84$ have a low average electron density and high metallicity, compared to galaxies in the field, with a larger presence of Hα emitters within filaments (Darvish et al. 2014, 2015). Thus, [O II] emitters, which are found in a large range of different environments (Hayashi et al. 2020), are also expected to be more common in filaments than in the field.

Detailed observations from MUSE suggest that filaments assist gas cooling, enhancing the star formation in galaxies (Vulcani et al. 2019). Hydrodynamical simulations have shown similar results (e.g. Liao & Gao 2019). These results could be a consequence of the outskirts of filaments being vorticity rich regions in simulations (e.g. Laigle et al. 2015), dominated by smooth accretion (e.g. Kraljic et al. 2019). Using cosmological hydrodynamic simulations (Cui et al. 2012; Cui, Borgani & Murante 2014), Cui et al. (2019) quantified that, at $z = 1$, ∼68 per cent of the gas is cold. About 48 per cent of this cold gas, $T < 10^5$ K, was found in sheets and 28 per cent in filaments. Furthermore, at $z = 1$ most haloes (∼70–100 per cent) with masses in the range $10^{11} \lesssim \mathrm{M_{halo}}(h^{-1}\,\mathrm{M_\odot}) \lesssim 3 \times 10^{13}$ live in filaments. Haloes in sheets are on average less massive than those found in filaments. The percentage of haloes in filaments with masses of $\sim 10^{10} h^{-1} M_\odot$ drops to ∼ 45 per cent with sheet haloes increase to ∼40 per cent (Cautun et al. 2014). As cold, dense gas that locates in haloes is needed for star formation to happen, we expect to find more model star-forming galaxies in filaments and sheets at $z = 1$.

Star-forming ELGs at $z \sim 1$ have been found to populate haloes of masses $\sim 10^{12} h^{-1}\,\mathrm{M_\odot}$ (Favole et al. 2016; Khostovan et al. 2018; Guo et al. 2018a) and have a linear bias around 1.5 (Comparat et al. 2013c; Guo et al. 2018a). Using our previous semi-analytical model (SAM) of galaxy formation (Gonzalez-Perez et al. 2018, hereafter GP18), we found model ELGs living in haloes with masses consistent with the observations but less clustered on large scales. As we found the percentage of model satellite ELGs to be below that in Favole et al. (2016), we argued that this could drive the differences found in the clustering of these objects. The fraction of model star-forming satellite galaxies is related to the modelling of the gas cycle (cooling, accretion, star formation, death of stars, etc.). Nevertheless, the lack of assembly bias in the models used to interpret the observations of star-forming ELGs might be partly responsible for this difference (Contreras et al. 2019).

Both observations and models show that star-forming galaxies in general, and ELGs in particular, populate dark matter haloes in a different way than mass selected samples (e.g. Zheng et al. 2005; Favole et al. 2016; Cochrane & Best 2018; Gonzalez-Perez et al. 2018; Guo et al. 2018a; Alam et al. 2019; Contreras et al. 2019).

Here, our objectives are to (i) characterize how ELGs trace the cosmic web and (ii) put the model [O II] emitters populations into context, by comparing their properties with those of star formation rate (SFR) and stellar mass selected samples. The star formation histories of galaxies are affected by their environment. In turn, this can have an impact on their relation with their host haloes and thus, their assembly bias. Our first objective, as stated above, is to start to understand these connections and their relevance both for modelling galaxy formation and in terms of their influence on the estimation of cosmological parameters. Our second objective is to contrast ELGs with other cosmological tracers. We expect the properties of star-forming ELGs to follow many of the trends found for the population of star-forming galaxies and thus, it will be useful to know when special care is needed to model these tracers, beyond assuming that a simple proxy, such as a cut in SFR, is adequate.

Here, we use the results from a SAM of galaxy formation and evolution. We use an updated version of the model presented in GP18. In Fig. 12 from that paper we showed that, qualitatively, model [O II] emitters trace filaments better than mass-selected galaxies. The distribution of matter on large scales can be segmented into four dynamically distinct environments: knots, filaments, sheets, and voids (e.g. Klypin & Shandarin 1983; Geller & Huchra 1989; Bond, Kofman & Pogosyan 1996). These four different cosmological structures are a natural outcome of gravitational collapse. Many methods have been developed to classify/identify these cosmological structures, and we refer the reader to Libeskind et al. (2018) for a detailed description. Here, we measure the large-scale environment of the dark matter simulation using two algorithms, one that uses a shear tensor, VWEB , and the other a tidal one, PWEB (Cui et al. 2018, 2019). Both methods use the eigenvalues of the Hessian matrix for the indicator field (velocity and potential, respectively) to spatially separate out these structures.

The plan of this paper is as follows. In Section 2.1, we introduce an updated version of the GALFORM SAM for galaxy formation and evolution. The VWEB and PWEB methods, used to characterize the large-scale environment, are described in Section 2.2 and in Appendix B. The selection of model [O II] emitters is presented in Section 3. The results on how [O II] emitters populate the cosmic web can be found in Section 3.2. Fixed number density samples are defined in Section 4, and their properties in different large-scale environment are presented in Section 4.2.2. In Section 5 we summarize our results.

2 METHODS

In this work, we analyse the $z = 0.83$ and $z = 1$ outputs from the MS-W7 N-body simulation (Guo et al. 2013; Gonzalez-Perez et al. 2014; Jiang et al. 2014). This N-body simulation is run within a box of 500 h^{-1} Mpc comoving side and assumes a cosmology consistent with the seventh year release from WMAP (Komatsu et al. 2011): matter density $\Omega_{m,0} = 0.272$, cosmological constant $\Omega_{\Lambda,0} = 0.728$, baryon density $\Omega_{b,0} = 0.0455$, a normalization of density fluctuations given by $\sigma_{8,0} = 0.810$, $n_s = 0.967$, and a Hubble constant today of $H(z = 0) = 100h\ \mathrm{km\,s^{-1}Mpc^{-1}}$ with $h = 0.704$.

The MS-W7 simulation has been populated with galaxies using a SAM for galaxy formation and evolution, based on that described in Gonzalez-Perez et al. (2018). This model is introduced in Section 2.1.

The large-scale environmentof the whole simulation box has been classified into knots, filaments, sheets, and voids using two algorithms: VWEB, which uses a velocity–shear–tensor, and PWEB , which uses a tidal–tensor (Cui et al. (2018, 2019). These algorithms are described in Section 2.2, and their resolution and threshold setting are further discussed in Appendix Section B.

2.1 The GALFORM galaxy model

SAMs use simple, physically motivated rules to follow the fate of baryons in a universe in which structure grows hierarchically through gravitational instability (see Baugh 2006; Benson 2010;

Table 1. Differences between the model presented here and the Gonzalez-Perez et al. (2018), GP18, GALFORM implementation. The ram-pressure stripping parameter $\epsilon_{\rm strip}$ is described in equation (6) of Font et al. (2008), and controls the stripping efficiency for reheated gas after the initial stripping event at the first pericenter (see also Benson & Bower 2010). In the model used here, the ram-pressure stripping happens on longer timescales than in GP18. Here, we use the updated model for the evolution of the super massive black holes (SMBH), as described in Griffin et al. (2019). This model assumes a mass of $10\,h^{-1}\,{\rm M}_{\odot}$ for the black hole seeds. The last three parameters in the table, $f_{\rm Edd}$, $\epsilon_{\rm heat}$, and $\alpha_{\rm cool}$, control the AGN feedback efficiency and their definitions can be found in section 3.5.3 of Lacey et al. (2016).

GALFORM parameter	GP18	This work
$\epsilon_{\rm strip}$	0.1	0.01
$M_{\rm BHseed}(h^{-1}\,{\rm M}_{\odot})$	0.	10.
$f_{\rm Edd}$	0.039	0.01
$\epsilon_{\rm heat}$	0.016	0.02
$\alpha_{\rm cool}$	0.9	0.8

Somerville & Davé 2015, for an overview of hierarchical galaxy formation models).

GALFORM was introduced by Cole et al. (2000) and since then it has been enhanced and improved (e.g. Baugh et al. 2005; Bower et al. 2006; Lagos et al. 2011; Lacey et al. 2016; Griffin et al. 2019). GALFORM follows the physical processes that shape the formation and evolution of galaxies, including (i) the collapse and merging of dark matter haloes; (ii) the shock heating and radiative cooling of gas inside dark matter haloes, leading to the formation of galaxy discs; (iii) quiescent star formation in galaxy discs that takes into account both the atomic and molecular components of the gas (Lagos et al. 2011); (iv) feedback from supernovae, from active galactic nuclei (Bower et al. 2006) and from photoionization of the intergalactic medium; (v) chemical enrichment of the stars and gas (assuming instantaneous recycling); (vi) galaxy mergers driven by dynamical friction within dark matter haloes. GALFORM predicts the number and properties of galaxies that reside within dark matter haloes of different masses and assembly histories. This information can be described in terms of a non-parametric halo occupation distribution (HOD) function, i.e. the mean number of galaxies as a function of halo mass.

Currently, there are two main branches of GALFORM: one with a universal stellar initial mass function (Gonzalez-Perez et al. 2014, 2018, and this work) and one that assumes different IMFs for quiescent and burst episodes of star formation (Lacey et al. 2016). These two models have also been re-calibrated to run on a dark matter simulation with a Planck cosmology (Planck Collaboration XVI 2014; Baugh et al. 2018).

Here, we have modified the SAM described in Gonzalez-Perez et al. (2018), hereafter GP18, to better match the observed passive fraction of galaxies at $z = 0$ and to include the updated treatment of the evolution of supermassive black holes (SMBHs) introduced by Griffin et al. (2019). These two aspects are described in more detail in Sections 2.1.1 and 2.1.2, respectively. The parameters that have been modified are summarized in Table 1, other free parameters have been inherited without change from the model described in GP18.

The free parameters presented in Table 1 have been calibrated against observations at $z = 0$: the luminosity function in the b_J and K bands (focusing on the region around the knee), the observed black hole–bulge mass relation (Fig. 1) and the local passive fraction (Fig. 2). When calibrating the model presented here, our aim was to make the smallest number of changes to the GP18 model parameters, despite the introduction of the updated scheme for black hole growth.

2.1.1 *The treatment of gas in satellite galaxies*

The model we present here has been modified to achieve a local passive fraction of galaxies that is closer to that inferred from observations. We use the model of Font et al. (2008) for the gradual ram-pressure stripping of hot halo gas from satellite galaxies. In this model, the reservoirs of hot halo gas and reheated gas associated with the galaxy when it becomes a satellite are partially stripped according to the results of hydrodynamical simulations (McCarthy et al. 2008). This mainly happens at the first pericenter of the satellite orbit, in the denser central parts of the hot halo where the ram pressure is highest. Ram-pressure stripping of the gas ejected from the galaxy by supernova feedback after it becomes a satellite is assumed to occur at an efficiency that is reduced by a factor $\epsilon_{\rm strip}$ compared to the initial stripping (as described in equation 6 from Font et al. 2008;

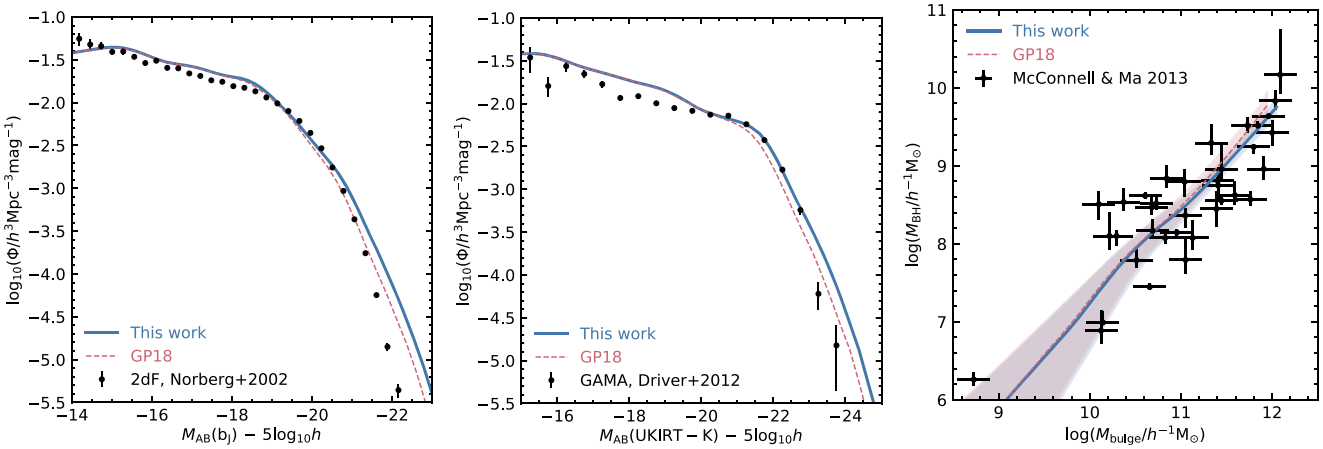

Figure 1. The predicted luminosity functions at $z = 0$, in the b_J band ($\lambda_{\rm eff} = 4500$ Å, left) and in the K band ($\lambda_{\rm eff} = 2.2$ µm, middle), compared with observations from Norberg et al. (2002) and Driver et al. (2012), respectively. The right-hand panel shows the super massive black hole mass versus bulge stellar mass relation at $z = 0$ compared to observational data from McConnell & Ma (2013). The lines in the right-hand panel show the median of the predicted super massive black hole mass in bins of bulge mass and the shading the 10–90 percentiles of the model distributions. The blue-dashed lines show the predictions from the GP18 model, while the red solid lines show the predictions from the model presented here. These data sets were used to calibrate the free parameters of the model.

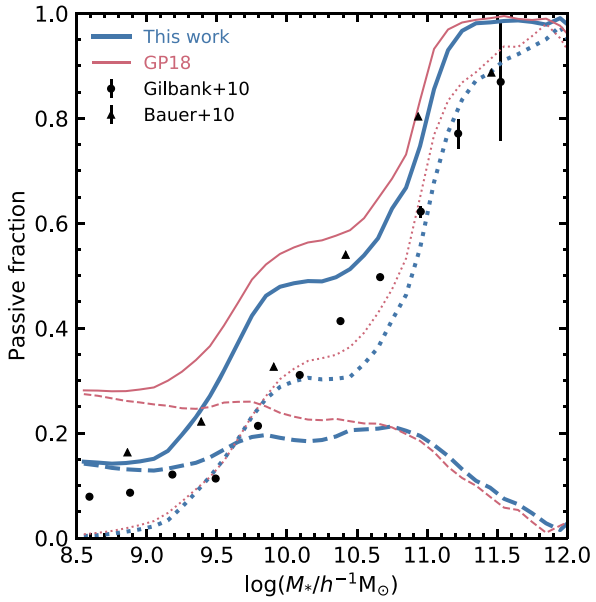

Figure 2. The fraction of passive galaxies at $z = 0$, i.e. those with sSFR $< 0.3/t_{\rm Hubble}(z = 0)$, for this work (the thick lines) and the GP18 model (the thin lines), compared to the observational results from Gilbank et al. (2010; the circles) and Bauer et al. (2013; the triangles) as extracted and presented in Furlong et al. (2015). The solid lines show the total passive fraction, the contribution from satellite galaxies is shown by the dashed lines, and that from centrals by the dotted lines.

see also Benson & Bower 2010). This extra suppression is invoked as the reheated gas is being ejected throughout the orbit of the satellite, not just at the pericenter. The satellite will spend more of its time in the outer parts of the hot halo where the ram pressure is lower, so less of the reheated gas should be stripped. In the model used here, we set $\epsilon_{\rm strip} = 0.01$, which is 10 times lower than the value of $\epsilon_{\rm strip} = 0.1$ assumed in Font et al. (2008) and in GP18, so that ram-pressure stripping after the first pericenter takes place 10 times more slowly than in GP18. The main effect of this change, as seen in Fig. 2, is to reduce the passive fraction for low-mass satellite galaxies in the model. In fact, the passive fraction at $z = 0$ is practically insensitive to the precise value of $\epsilon_{\rm strip}$ once it is around 0.01 or lower. Quiescent star formation in a satellite is followed by supernovae feedback that heats up the cold gas. This gas needs to cool down again before being available for further star formation and so a minimum passive fraction is reached for small satellite galaxies.

The separation of galaxies into passive and star-forming is done using the specific star formation rate (sSFR; i.e. the ratio between SFR and stellar mass) boundary proposed by Franx et al. (2008): $sSFR = 0.3/t_{\rm Hubble}(z)$. Fig. 2 shows that a slower removal of hot gas in satellites reduces the number of passive satellite galaxies. These dominate the low-mass end. Around masses $M_* \sim 10^{10}\,h^{-1}$M$_\odot$ at $z = 0$, the model passive fraction displays a plateau dominated by the contribution of discs. This region of model masses is sensitive to the efficiency of the model AGN feedback. A more detailed exploration of the impact of the treatment of gas in models of galaxy formation will be developed in a future paper.

2.1.2 *The growth of model supermassive black holes*

Griffin et al. (2019) presented a new version of GALFORM with an updated model of the growth and spin evolution for SMBHs. This model assumes that SMBHs can grow in three different ways: (i) during starbursts triggered either by galaxy mergers or disc instabilities, (ii) by accreting gas from the hot atmosphere of massive haloes, and (iii) by SMBH–SMBH mergers after a galaxy merger. The model takes into account how the angular momentum of both the SMBH and the accretion disc affects the consumption of gas.

In this updated model, SMBHs grow from seeds with mass $M_{\rm seed}$. When a galaxy is formed in the model, it is assigned a black hole of mass $M_{\rm seed}$. The value of this $M_{\rm seed}$ is a free parameter set, which we set to 10 h^{-1}M$_\odot$ (see also Table 1). For $M_{\rm seed} > 0$, the black hole properties for SMBH in the observed mass range converge rapidly (for further discussion see Griffin et al. 2019).

The angular momentum of the gas in the inner accretion disc is assumed to be periodically randomized with respect to the angular momentum of the SMBH, i.e. we assume a 'chaotic accretion' mode (King, Pringle & Hofmann 2008; Fanidakis et al. 2011).

This updated model includes the evolution of SMBH spins. This evolution affects the growth of SMBHs and therefore the AGN activity. Griffin et al. (2019) showed the new predicted AGN luminosity functions for a range of wavelengths. The SMBH mass versus bulge mass relation at $z = 0$ is shown in the right-hand panel of Fig. 1. The updated model has a distribution consistent with that in the GP18 model.

2.1.3 *The emission line model*

Nebular emission lines are produced by gas heated by newly formed stars or nuclear activity. Here, we model the star-forming contribution. In GALFORM, the ratio between the [O II] luminosity and the number of Lyman continuum photons is calculated using H II region models of Stasińska (1990). The GALFORM model uses H II region models tabulated for a range of gas metallicities but with a uniform density of 10 hydrogen particles per cm^{-3} and one ionizing star in the centre of the region with an effective temperature of 45 000 K. The ionization parameter of the H II region models is around 10^{-3}, with the exact value depending on the metallicity in a non-trivial way. These ionization parameters are averages over the grid of H II regions provided by Stasińska (1990). Further details on the emission line model can be found in GP18 (see also Orsi et al. 2008).

Nebular emission lines are assumed to be attenuated by dust in the same way as the stellar continuum (Gonzalez-Perez et al. 2013; Lacey et al. 2016).

The properties of model ELGs derived using the Stasińska (1990) default models are consistent with those obtained using the Anders & Fritze-v. Alvensleben (2003) model for typical H II regions. The model emission line luminosity functions are also in reasonable agreement with the results derived from a model that assumes a large range of H II regions (Comparat et al. 2015). The nebular emission luminosity functions at different redshifts derived from this emission line model were found to be in agreement with observations (Lagos et al. 2014; Gonzalez-Perez et al. 2018).

2.2 The cosmic web

Here, we apply two algorithms, VWEB and PWEB , to classify the large-scale environmentof the whole simulation box into knots, filaments, sheets, and voids (Cui et al. 2018, 2019).

The VWEB method uses a dimensionless velocity–shear–tensor as the tracer to classify the large-scale environment. Following Hoffman et al. (2012), at a given redshift, z, the velocity, $\vec{v}(\vec{r})$, shear tensor is

Table 2. The cuts applied to the model galaxies in order to mimic the selection of [O II] emitters in the corresponding observational survey are the same as those summarized in table 2 from Gonzalez-Perez et al. (2018), except for the eBOSS-SGC survey. We apply here the colour cuts described in Raichoor et al. (2017) for the eBOSS-SGC selection (further details can be found in Appendix A) plus a cut in [O II] flux to mimic the instrumentation limitation of the eBOSS-SGC survey. The magnitudes are on the AB system. The particular filter response used for the different cuts is indicated by a superscript on the magnitude column.

Cuts to mimic	Apparent magnitude	[O II] flux ($\mathrm{erg\,s^{-1}cm^{-2}}$)	Colour selection
DEEP2	$R_{\rm AB}^{\rm DEIMOS} < 24.1$	2.7×10^{-17}	None
VVDS-Deep	$i_{\rm AB}^{\rm CFHT} \leq 24$	1.9×10^{-17}	None
VVDS-Wide	$i_{\rm AB}^{\rm CFHT} \leq 22.5$	3.5×10^{-17}	None
eBOSS-SGC	$21.825 < g_{\rm AB}^{\rm DECam} < 22.825$	1×10^{-16}	$-0.068(r-z)+0.457 < (g-r) < 0.112(r-z)+0.773$ and $0.218(g-r)+0.571 < (r-z) < -0.555(g-r)+1.901$ $(r-z) > 0.3$ and $(g-r) > -0.3$ and $(g-r) < 1.1 \cdot (r-z) - 0.13$ and $(g-r) < -1.18 \cdot (r-z) + 1.6$
DESI	$r_{\rm AB}^{\rm DECam} < 23.4$	8×10^{-17}	

defined as

$$\Sigma_{\alpha\beta} = -\frac{1}{2H(z)}\left(\frac{\partial v_\alpha}{\partial r_\beta} + \frac{\partial v_\beta}{\partial r_\alpha}\right), \quad (1)$$

where, $H(z)$ is the Hubble constant at redshift z. The eigenvalues of $\Sigma_{\alpha\beta}$ are denoted as λ_i^V ($i = 1$, 2, and 3).

The PWEB method classifies the large-scale environment based on the tidal tensor, which is measured with the Hessian matrix of the gravitational potential field, $\phi(\vec{r})$. The gravitational potential is calculated from the matter density distribution via the Poisson equation, $\nabla^2\phi = 4\pi G\rho\delta$, where ρ is the mean density and δ is the density fluctuation. The tidal tensor, with units s^{-2}, is defined as follows (Hahn et al. 2007):

$$P_{\alpha\beta} = \frac{\partial^2\phi}{\partial r_\alpha \partial r_\beta}. \quad (2)$$

The computation of the eigenvalues for both matrices is performed on the particles of the MS-W7 dark-matter-only simulation, split in regular 512^3 cells grids. The typical side of a grid cell is $\sim 1\,h^{-1}$ Mpc. We use a triangular-shaped cloud in cell prescription, for obtaining a smoothed density and velocity distribution at each point on the grid. These are smoothed further over a scale of $\sim 5\,h^{-1}$ Mpc. Then, for every grid cell, the eigenvalues of the velocity–shear–tensor, tidal–tensor for PWEB , are computed according to equation (1) for VWEB and equation (2) for PWEB . Note that although neither VWEB nor PWEB directly use the dark matter particles for their calculations, the larger the number of particles per cell, the more accurate the velocity or potential fields will be. The large smoothing scale, $\sim 5\,h^{-1}$ Mpc, provides a robust velocity and density field in real space, and thereby reliable tensors.

Each individual cell is then classified as either 'void', 'sheet', 'filament', or 'knot' according to the eigenvalues $\lambda_1 > \lambda_2 > \lambda_3$:

(1) Void, if $\lambda_1 < \lambda_{\rm th}$,
(2) Sheet, if $\lambda_1 \geq \lambda_{\rm th} > \lambda_2$,
(3) Filament, if $\lambda_2 \geq \lambda_{\rm th} > \lambda_3$,
(4) Knot, if $\lambda_3 \geq \lambda_{\rm th}$,

where $\lambda_{\rm th}$ is a free threshold parameter (Hoffman et al. 2012; Libeskind et al. 2012, 2013). Following Carlesi et al. (2014) and Cui et al. (2018), we find that the threshold adopted for VWEB at $z = 0$, $\lambda_{\rm th}^V = 0.1$, is also suitable for the simulation results at higher redshift (see Cui et al. 2019, for the redshift evolution of the mass and volume fractions of these large-scale structures with this fixed threshold). This value gives very convincing structures, which are visually comparable with the density field (see Appendix B for further details). However, we find that the threshold $\lambda_{\rm th}^P = 0.01 s^{-2}$ at $z = 0$ for the PWEB method results in slightly larger voids and sheet structures at the two redshifts investigated in this paper, $z = 0.83$, 0.99. Therefore, we lower the threshold to $0.005\,\mathrm{s}^{-2}$ for PWEB to provide consistent structures to VWEB .

The effects of resolution and the choice of different thresholds for the cosmic web classification are discussed in detail in the Appendix B.

3 MODEL ELGS

A range of cosmological surveys, such as SDSS/eBOSS-SGC (Dawson et al. 2016) and DESI (DESI Collaboration 2016), are or will be targeting star-forming ELG galaxies to probe the nature of dark energy using spectroscopic redshifts. ELGs selected with optical instruments at $z \sim 1$ are dominated by [O II] emitters (Comparat et al. 2015). As we previously did in GP18, we select [O II] emitters from the SAM described in Section 2.1 mimicking the samples from different surveys, Section 3.1. We then explore how these model galaxies trace the large-scale environmentin Section 3.2.

In this work, we use a model that strips the gas in satellite galaxies slower than in GP18. This modification has a strong impact on decreasing the passive fraction of galaxies with masses below $10^{10}h^{-1}\,\mathrm{M}_\odot$, however, this merely changes the fraction of model [O II] emitters by up to 5 per cent. Model ELGs at $z \sim 1$ are mostly centrals, with a satellite fraction between 4 per cent and 9 per cent, dominated by star-forming galaxies with $sSFR > 0.3/\mathrm{t}_{\rm Hubble}(z)$.

3.1 Model ELGs sample selection

We select model ELGs using the cuts specified in Table 2 in apparent magnitude, [O II] flux, and colour. The magnitude and flux cuts reproduce the limits in the DEEP2 (Newman et al. 2013) and VVDS (Le Fèvre et al. 2013) surveys, applied to select the corresponding model [O II] emitters. No further colour cuts are applied to the model DEEP2 and VVDS selections, as the observational colour cuts were applied to restrict the redshift range and here we are limiting our study to two single simulation outputs at $z = 0.83$ and $z = 0.99$. We have additional colour cuts to select model DESI (DESI Collaboration 2016) and SDSS-IV/eBOSS (Raichoor et al. 2017) ELGs. These colour cuts were set observationally to target an spectroscopic galaxy sample with colours that minimally overlap with those from stars (further details can be found in appendix A).

Here we focus on two simulation outputs at $z = 0.83$ and $z = 0.99$, which are separated by 717 Myr. The lowest of these two redshifts

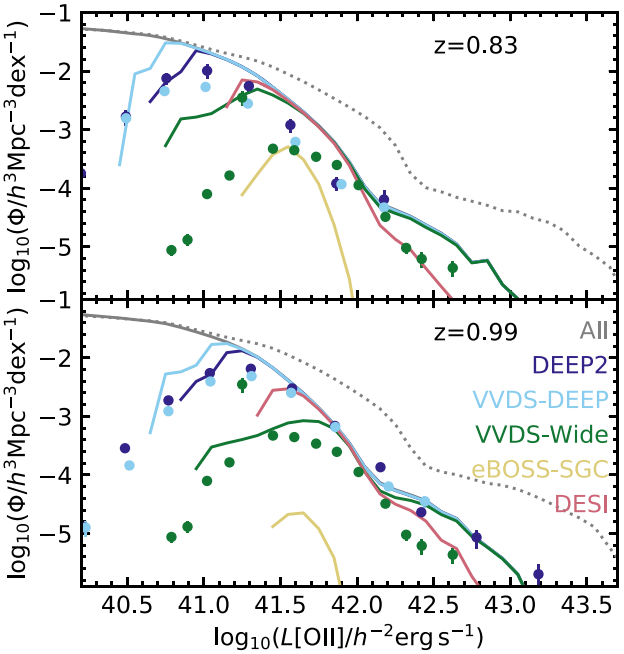

Figure 3. The luminosity function of [O II] emittersat $z = 0.83$ and $z = 0.99$ for model galaxies selected with the DEEP2 (the dark blue lines, mostly overplotted), VVDS-DEEP (the light blue lines), VVDS-Wide (the green lines), eBOSS-SGC (the yellow lines), and DESI (the red lines) cuts given in Table 2. The grey solid lines the model total dust attenuated L[O II] luminosity function, the intrinsic one is shown by the dotted grey lines. The data from DEEP2 and VVDS are colour coded like the model galaxies selected to mimic both surveys. The observational errors come from jackknife re-sampling (Comparat et al. 2016a) and in some cases are smaller than the corresponding symbol.

Table 3. Fraction of ELGs in the different large-scale environment structures as classified by the VWEB algorithm. The percentage of satellite galaxies for each selection is shown in brackets.

$z = 0.83$	VVDS-DEEP (11 per cent)	eBOSS-SGC (5 per cent)
Voids	0.05 (3 per cent)	0.04 (1 per cent)
Sheets	0.34 (6 per cent)	0.32 (3 per cent)
Filaments	0.48 (11 per cent)	0.51 (5 per cent)
Knots	0.13 (24 per cent)	0.12 (13 per cent)
$z = 0.99$	DEEP2 (7 per cent)	DESI (4 per cent)
Voids	0.04 (2 per cent)	0.04 (2 per cent)
Sheets	0.32 (4 per cent)	0.34 (2 per cent)
Filaments	0.51 (7 per cent)	0.51 (4 per cent)
Knots	0.13 (15 per cent)	0.12 (8 per cent)

is close to the effective redshift of the SDSS/eBOSS-SGC sample (Raichoor et al. 2017), $z = 0.84$, which in turn is close to the average redshift of VVDS-DEEP (Comparat et al. 2015). The VVDS-Wide sample has a lower average redshift and will not be included in the clustering and environment analysis presented later in this work. The DESI ELG sample is designed to have a redshift baseline between 0.6 and 1.7 and an anticipated effective redshift of $z \sim 1$ (DESI Collaboration 2016). The clean sample of DEEP2 ELGs has a mean redshift of 0.97 (Comparat et al. 2017). Both values are close to the redshift $z = 0.99$ of the simulation output.

3.1.1 *[O II] luminosity function*

Model [O II] emitters are selected in numbers that are in reasonable agreement with observational selections, as shown in Fig. 3. Note that the dust attenuation in this model is such, that it mostly affects the most luminous and massive [O II] emitters. As reported in GP18, the change in slope of the luminosity functions shown in Fig. 3 is due to galaxies with an ongoing starburst that dominate the bright end, $L[\mathrm{OII}] > 10^{42} h^{-2} \mathrm{erg\,s^{-1}}$. This bright end is also dominated by galaxies with a bulge to total mass above 0.5 (spheroids) and compact, with half-mass radii smaller than $0.5 h^{-1}$ kpc. The luminosity functions shown in Fig. 3 are similar to those in GP18 and accompany Errata.[11]

The number density of model SDSS/eBOSS-SGC ELGs at $z \sim 0.83$, $\sim 1.58 \times 10^{-4} h^3 \mathrm{Mpc}^{-3}$, is below the current observational estimations $\sim 2.67 \times 10^{-4} h^3\,\mathrm{Mpc}^{-3}$ (Raichoor et al. 2017). In Guo et al. (2019), it was presented an empirical model directly calibrated with SDSS/eBOSS-SGC data and they compared their results with the model galaxies presented here. From this comparison, it appeared that besides lacking satellite ELGs, as it was concluded in GP18, there might be a lack of massive central galaxies. This is also suggested by the results presented in Comparat et al. (2017) for DEEP2. Although dust attenuation affects the most luminous and massive galaxies, there might be other physical processes contributing to the discrepancies found, from the simplicity of our emission line modelling to a more fundamental aspect of the growth of massive galaxies (Mitchell et al. 2018).

3.2 Model ELGs in the cosmic web

The large-scale environment of the dark matter in the N-body simulation has been classified using the algorithms described in Section 2 into voids, sheets, filaments, and knots. Table 3 summarizes how model ELGs are distributed within the different structures of the cosmic web, as classified by the VWEB algorithm, although similar results are found when using PWEB. We find that about 80 per cent of model ELGs live in either filaments or sheets, with half of them in filaments.

The distribution of ELGs in the cosmic web, summarized in Table 3, is also reflected in the split of the [O II] luminosity function. This is shown in Fig. 4, for DESI model galaxies at $z = 0.99$, classified using the VWEB algorithm (similar results are found for PWEB). The [O II] luminosity function varies in normalization for the different large-scale environment structures, but the shape changes minimally. The brightest model [O II] emitters are found in the structures where they are most dominant: filaments and sheets.

As the [O II] luminosity function, the SFR function also shows different normalisations but similar shapes for galaxies in different cosmic web structures. Fig. 5 shows the case for DESI model galaxies at $z = 0.99$ classified with VWEB . Note that in GALFORM all galaxies have an SFR above zero, even if very small in some cases.

Fig. 5 also shows the distribution of model galaxies in the SFR–stellar mass plane. It is clear from here, that model ELGs are not directly equivalent to imposing a cut in SFR. This was also reported in GP18 and is common to all the studied ELG selections.

The galaxy stellar mass function for DESI model galaxies is also shown in Fig. 5. In this case, there is a clear change in the shape for galaxies in knots, at high masses. Model ELGs in knots tend to be

[11] Due to a problem with filter naming, the selection for VVDS was effectively done with the *r* band, instead of the indicated *i* band. This discrepancy has been corrected in this work.

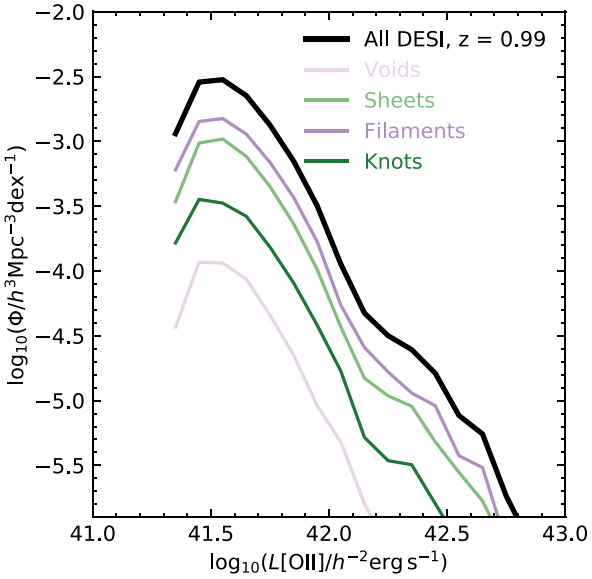

Figure 4. The [O II] luminosity function for the model DESI model galaxies at $z = 0.99$, the thick line, and the contribution of the different large-scale environment structures, the thin lines, as classified by the VWEB algorithm (see the legend).

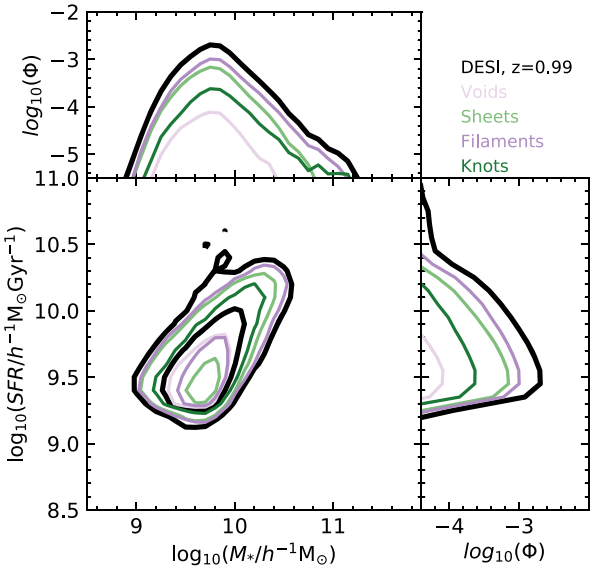

Figure 5. The $z = 0.99$ distribution of galaxies in the SFR-stellar mass plane for all DESI model galaxies, the thick lines, and those living in different large-scale environment structures, the thin lines as classified by the VWEB algorithm. The sSFR–stellar mass plane has been collapsed into the galaxy stellar mass function, top subpanel, and the SFR function, right subpanel. The corresponding densities shown are $\Phi(h^3\,\mathrm{Mpc}^{-3}\,\mathrm{dex}^{-1})$.

more massive. This is also found for the other ELG selections. This might be related with the larger fraction of satellite galaxies found in knots, as summarized in Table 3.

As knots appear in denser regions, haloes are expected to be more massive and, thus, able to host several galaxies. Given an ELG selection, knots tend to host more satellites, with the differences being the largest for satellite galaxies with stellar masses around $10^{10}h^{-1}\,\mathrm{M}_\odot$. This is the value above which the galaxy stellar mass

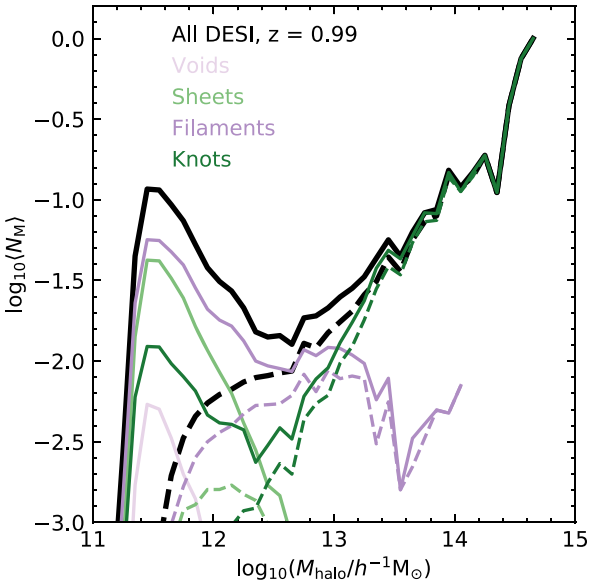

Figure 6. The mean halo occupation distribution (HOD) as a function of host halo mass for the model DESI galaxies at $z = 0.99$, the thick line, and the contribution of the different large-scale environment structures, the thin lines, as classified by the VWEB algorithm (see the legend). The solid lines show the total mean HOD; the contribution from satellite galaxies is shown as the dashed lines.

function in knots starts to differ from the the other cosmic web structures, as it can be seen in Fig. 5.

We also find that the percentage of satellite galaxies as a function of stellar mass varies significantly between the different ELG selections (not shown in figure).

In knots, the gas fuelling star formation in satellite galaxies will be removed after some time and little new gas will be fuel to those galaxies. This gas will feed the central galaxy. We defer to the future studying the evolution of the star formation in model galaxies populating different cosmic web structures.

3.2.1 The mean halo occupation distribution of ELGs

Fig. 6 shows the mean HOD for the DESI model galaxies at $z = 0.99$. This HOD is well below having one galaxy per halo. The HOD of model central ELGs is close to an asymmetric Gaussian with maybe a plateau (see also GP18). Galaxy mock catalogues from HOD models usually assume a very different shape from that seen in Fig. 6. The shape usually assumed for HOD models is that characteristic for stellar-mass-selected samples, this will be further explored in Section 4.

Fig. 6 shows that the normalization of the central galaxies peak decreases for different large-scale environment structures, following the trend in density reported in Table 3. The minimum halo mass to host an ELG remains practically independent of the cosmic web, except for voids, for which there is a slight increase in mass. The number of ELGs in voids is quite low, and those are mostly central. Note that the minimum halo mass in the model HOD shown in Fig. 6 is not affected by resolution effects.

In voids and sheets, there are almost no satellite ELGs. This can be seen in Fig. 6 for model DESI galaxies. The contribution of satellite galaxies is so small in voids and sheets that the global shape of the HOD for these environments can be described as an asymmetric Gaussian.

The shape of the mean HOD does change with environment. The HOD for central galaxies has a plateau in filaments and knots. There is a clear increase in the power law followed by model satellite ELGs in knots. The differences among the cosmic web structures highlight the importance that environmental processes have in shaping the evolution of galaxies. Environmental processes will therefore impact the small-scale clustering derived for galaxies populating different large-scale environments.

3.2.2 *The clustering of ELGs*

Here, we study the clustering of ELGs living in different large-scale environment. Model galaxies come from a periodic simulated box and therefore the configuration space two-point autocorrelation function, ξ, is calculated using a simple estimator: $1 + \xi(r) = 2DD/(n^2 V \mathrm{d}V)$, where DD is the number of distinct model galaxy pairs with separation between r and $r + \mathrm{d}r$ and the denominator is the average number of neighbours found in the volume $\mathrm{d}V$ of a spherical shell of radius r and thickness dr (see also Gonzalez-Perez et al. 2011). The dark matter two-point correlation function is calculated using the particles from the MS-W7 dark matter only simulation. The calculation of the two-point autocorrelation functions has been done using the publicly available code CUTE .[12]

The 3D pair counts, $DD_{\rm ccf}$, needed to estimate the cross-correlation, $\xi_{\rm All,\,X}$, between the whole ELG sample and the subsamples populating different large-scale environment structures is obtained with the publicly available PYTHON package `Corrfunc.theory.DD`.[13] The two-point cross-correlation is then estimated as $1 + \xi_{\rm All,\,X}(r) = DD_{\rm ccf}/(N_{\rm All} N_X \mathrm{d}V/V)$, where $N_{\rm All}$ is the number of all ELGs and N_X those ELGs in a given large-scale environment structure, within the simulation volume, V. The Poisson errors for the cross-correlation are estimated as $(1 + \xi_{\rm All,X})/\sqrt{(DD_X)}$, where DD_X is the number of unique pairs of the subsample of ELGs, living in either voids, sheets, filaments, or knots.

Fig. 7 shows the real-space two point autocorrelation function for all DESI model galaxies at $z = 0.99$ compared with that for the dark matter. At large scales, $r > 1h^{-1}$Mpc, DESI ELGs trace the dark matter clustering, with a linear bias close to 1. Table 4 presents the large-scale bias for each of the ELG samples studied here (see Table 2), which are all close to 1.

The properties of the model galaxies explored here are naturally affected by assembly bias, i.e. the dependence on halo assembly history as well as on halo mass (e.g. Zentner, Hearin & van den Bosch 2014). Therefore, the bias that we measure is the combination of the cosmological halo bias and that resulting from the assembly bias. The latter is expected to have a small or negative effect for star-forming samples, such as the ELGs we study here (Contreras et al. 2019). This implies that the bias measured from a catalogue of galaxies constructed with an HOD model without considering assembly bias, might be larger than the values reported here. To quantify the assembly bias, we present in Fig. 7 the clustering for the model DESI sample shuffled within haloes of similar mass (see e.g. Jiménez et al. 2019). In this case, we found a small but negative signal. For the other ELG selections, we find almost no galaxy assembly bias signal.

Fig. 7 also shows the cross-correlation between all DESI ELGs at $z = 0.99$ and those living in large-scale environmentas classified by the VWEB algorithm. At small scales, $r \leq 0.5h^{-1}$Mpc, the clustering of ELGs follows closely the autocorrelation function in all large-scale environment, except for knots from the DESI sample. This is clearer for the DEEP2 and VVDS-DEEP samples (not shown here), for which pairs of galaxies are found at separations smaller than $0.03h^{-1}$Mpc, extending the one-halo term clustering to smaller scales than the DESI one. At larger scales, $r \geq 0.5h^{-1}$ Mpc, differences are found for the clustering of ELGs living in different large-scale environment. ELGs in filaments are the ones clustered most similar to the two-point autocorrelation function.

ELGs living in knots are the most clustered for separations $1 \leq r(h^{-1}\,\mathrm{Mpc}) \leq 10$. This can be seen in Fig. 7 for DESI model galaxies, but it is also the case for the other ELG selections. The fraction of

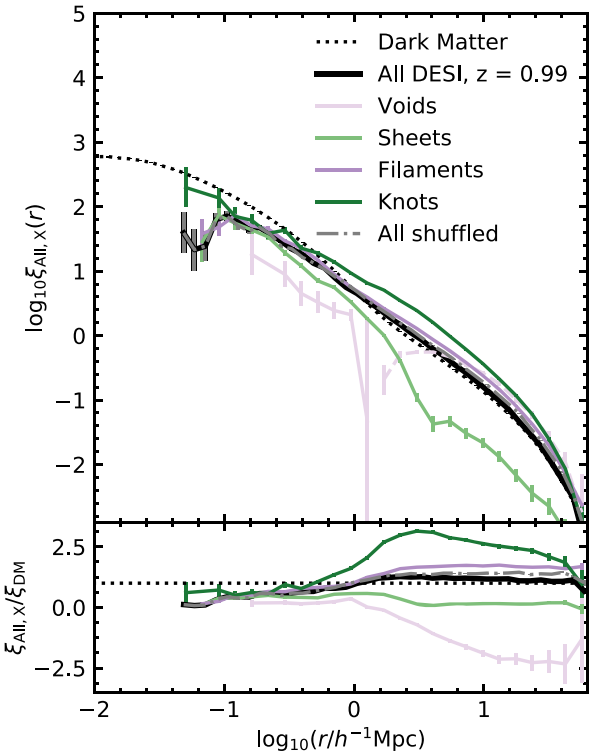

Figure 7. The real-space two-point correlation function, top panel, and ratio between the correlation functions of galaxies and of dark matter, $\xi_{\rm All,\,X}/\xi_{\rm DM}$. The dark matter autocorrelation function is shown as a dotted line. The autocorrelation function for all DESI model galaxies at $z = 0.99$ is shown as the thick solid lines and that for a shuffled sample as the dot–dashed lines. The thin solid lines show the cross-correlation, $\xi_{\rm All,\,X}$, between the whole ELG sample and ELGs in different large-scale environment structures as classified by the VWEB algorithm (see the legend). The logarithm of the absolute value of negative $\xi_{\rm All,\,X}$ is shown with the dashed thin lines. Poisson error bars are shown in both panels.

Table 4. Large-scale bias for the ELG samples presented in Table 2. The bias, b, and associated error have been obtained by minimizing χ^2 for the model galaxy autocorrelation function in real space: $\xi_{\rm gg} = b^2 \xi_{\rm Dark\ matter}$, in the range $8 \leq r(h^{-1}\,\mathrm{Mpc}) \leq 50$. This range comprises the large scales available for the underlying dark matter simulation, for which we can consider to be measuring the linear scale bias.

$z = 0.83$	VVDS-DEEP	eBOSS-SGC
	1.12 ± 0.01	1.09 ± 0.01
$z = 0.99$	DEEP2	DESI
	1.17 ± 0.01	1.07 ± 0.01

[12] https://github.com/damonge/CUTE (Alonso 2012).
[13] https://github.com/manodeep/Corrfunc (Sinha & Garrison 2020).

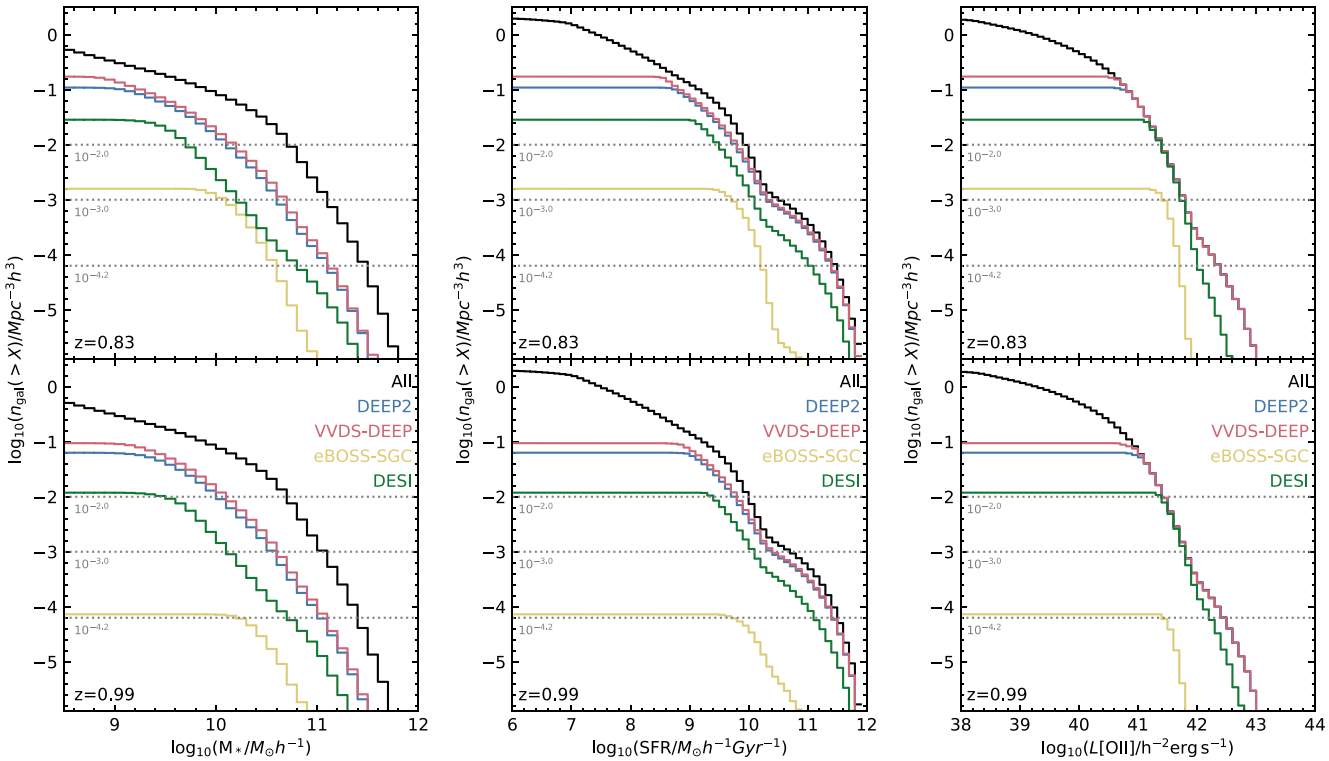

Figure 8. The cumulative abundance of all galaxies, the black lines, and ELGs selected as summarized in Table 2 and indicated in the legend, ranked by stellar L[O II] , left, SFR, middle, and stellar mass, right, at $z = 0.83$, top panel, and $z = 0.99$, bottom panel. The three number density cuts used to define similar samples, $n_{\rm gal} = 10^{-2}$, 10^{-3}, $10^{-4.2} h^3$ Mpc^{-3}, are indicated by the dashed horizontal lines.

satellite ELGs in knots is the largest found for the explored cosmic web structures (see Table 3). The large number of satellites might explain the reported boost in the clustering at intermediate scales, corresponding to the transition between the one-halo and two-halo terms.

For all the studied ELG samples, at large-scales, $r > 1h^{-1}$ Mpc, galaxies living in voids are less clustered than in any other environment, with cross-correlations having negative values and thus, negative bias values. There are few model ELGs found in voids (see Table 3), which are also the least dense regions, producing a low clustering at large scales. We find the differences to be dominated by the number density of each environment.

Similar results to those described above for the VWEB classification are found when using the PWEB algorithm to classify the large-scale environment.

4 ELGS IN CONTEXT

We have previously quantified how model ELGs trace the large-scale environment. ELGs are expected to typically trace less dense regions than mass-selected galaxies, such as luminous red galaxies (e.g. Alam et al. 2019). Here, we aim to contrast the properties of ELG samples with those selected using simpler criteria, such as stellar mass or SFR. This comparison will allow us to gain insight into which aspects of the ELG populations are unique and might be a source of additional systematic errors when they are used as cosmological tracers (see e.g. Avila et al. 2020).

To make a fair comparison between galaxy samples selected in different ways, we generate fixed number density samples with $n_{\rm gal} = 10^{-2}$, 10^{-3}, $10^{-4.2} h^3$ Mpc^{-3}. These fixed number density samples are generated by either imposing a single cut in stellar mass, SFR or L[O II] or starting with the ELG samples described in Section 3.1 and then imposing an extra cut in one of the three mentioned properties.

Given either the effective or mean redshifts of the different surveys considered in this study, as described in Section 3, the analysis is done at $z = 0.83$ for the eBOSS-SGC and VVDS-DEEP samples and at $z = 0.99$ for the DEEP2 and DESI ones.

4.1 Fixed number density samples

Fig. 8 presents the cumulative abundance of the whole galaxy population and the ELGs subsamples ranked by their stellar mass, SFR, and L[O II] . From here, making cuts in these three properties, fixed number density samples are constructed with $n_{\rm gal} = 10^{-2}$, 10^{-3}, $10^{-4.2} h^3$ Mpc^{-3}, for the all model galaxies and the four ELG selections.

Fig. 9 shows the SFR–stellar mass plane, galaxy stellar mass function and SFR function for samples with three different number densities selected either imposing cuts on the stellar mass, the SFR, or L[O II] on either the global population or the ELGs. Fig. 9 only shows the results for the model VVDS-DEEP sample, but similar trends are found for the other ELG selections studied here, which are summarized in Table 2. Galaxies selected by their SFR have stellar masses spreading a large dynamical range. This is also the case for galaxies selected with a cut in L[O II]. However, in this case, galaxies tend to have lower masses and SFR than the fixed number density sample selected with cuts in the SFR. Fig. 9 shows that, as reported in GP18, the model ELG selection is not equivalent to imposing a cut in SFR.

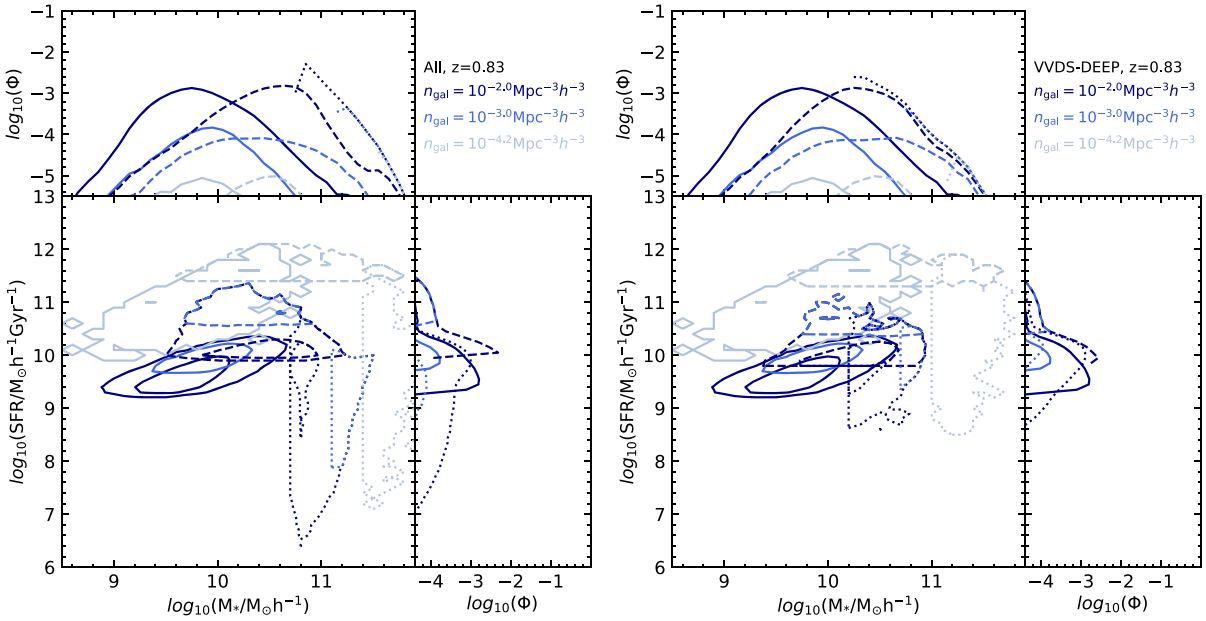

Figure 9. The $z = 0.83$ distribution of galaxies in the SFR-stellar mass plane for galaxies with the number densities indicated in the legend, selected by either imposing a single cut to the whole population of model galaxies (the left-hand panel) or an extra cut to the ELGs samples summarized in Table 2 (the right-hand panel shows VVDS-DEEP galaxies at $z = 0.83$). The properties used for selecting the fixed number density samples are stellar mass (the dotted lines), SFR (the dashed lines), and L[O II] (the solid lines). The sSFR–stellar mass plane has been collapsed into the galaxy stellar mass function, top subpanels, and the SFR function, right subpanels. The corresponding densities shown are $\Phi(h^3\,\mathrm{Mpc}^{-3}\,\mathrm{dex}^{-1})$.

The median mass of the host haloes increases with stellar mass limit (not shown), which decreases the number density of the galaxy sample. This trend is not found as clearly when either the SFR or L[O II] are used to select the fixed number density samples. In this case, the selected galaxies spread a large range of stellar masses even for low number densities, as can be seen in Fig. 9.

The median L[O II] for fixed number density galaxy samples selected with a single cut in their stellar mass is below $10^{39.5}h^{-2}\mathrm{erg\,s^{-1}}$, while all the ELG fixed number density selections and those made with a single cut in SFR and L[O II] have median L[O II] above $10^{40.5}h^{-2}\mathrm{erg\,s^{-1}}$. Fixed number density ELGs have SFR $> 10^8 h^{-1}\,\mathrm{M_\odot\,Gyr^{-1}}$ at the studied redshifts. The median SFR increases with decreasing number density when galaxies are selected by their SFR and their L[O II] .

4.1.1 Mean HOD of fixed number density samples

The mean HOD for fixed number density central galaxies selected with a cut in their stellar mass follows a soft step function, reaching unity: at least one galaxy of a given mass will be found in large enough haloes (see the top panel in Fig. 10). This is very different from the behaviour of fixed number density SFR-selected central galaxies (Zheng et al. 2005; Geach et al. 2012; Contreras et al. 2013; Cochrane et al. 2017, 2018). As shown in Fig. 10, these follow a shape closer to an asymmetric Gaussian plus a shallow power law (Cochrane & Best 2018; Gonzalez-Perez et al. 2018). A star-forming galaxy is not found in all haloes above a certain mass. The top panel in Fig. 10 shows that the same is true for fixed number density galaxies selected with a single cut in L[O II]. In this case, the suppression in the number of central galaxies found in massive haloes is even larger than for fixed number density samples selected with an SFR cut.

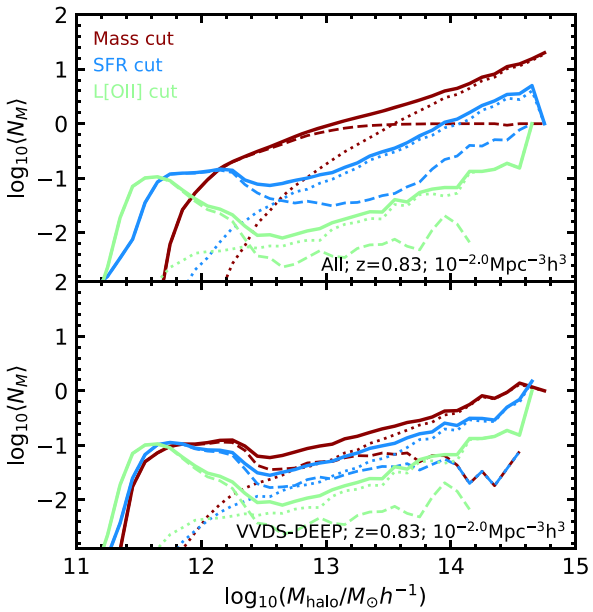

Figure 10. The $z = 0.83$ mean halo occupation distribution (HOD) for fixed number density samples, $10^{-2}h^{-3}\mathrm{Mpc^{-3}}$, of galaxies selected with either a single cut in one of the properties specified in the legend (top panel) or applying a cut in one of those properties to the VVDS-DEEP (bottom panel). The dark red lines show the HOD when the number density selection is done using a cut in stellar mass, the blue lines when this is done using a cut in SFR and the green lines when a cut in L[O II] is applied. The solid lines correspond to the total HOD, while the contribution from centrals is shown with the dashed lines and that of satellites by the dotted lines.

Table 5. Bias for the fixed number density samples described in Section 4, at $z = 0.83$ and $z = 0.99$. The bias is tabulated for the stellar mass, SFR, and L[O II]-selected samples separated by a comma. The bias and associated error have been obtained as described in Section 3.2.2 in the range $8 \leq r(h^{-1}\mathrm{Mpc}) \leq 50$.

z	Survey	$10^{-2}h^3\,\mathrm{Mpc}^{-3}$	$10^{-3}h^3\,\mathrm{Mpc}^{-3}$	$10^{-4.2}h^3\,\mathrm{Mpc}^{-3}$
0.83	All	1.90 ± 0.01, 1.27 ± 0.01, 0.99 ± 0.01	2.56 ± 0.03, 1.21 ± 0.03, 0.97 ± 0.05	4.36 ± 0.06, 1.48 ± 0.68, 1.76 ± 0.96
	VVDS-DEEP	1.26 ± 0.01, 1.13 ± 0.01, 0.99 ± 0.01	1.50 ± 0.04, 1.18 ± 0.02, 0.96 ± 0.09	2.35 ± 0.40, 1.11 ± 0.71, 1.83 ± 0.91
	eBOSS-SGC	-, -, -	1.23 ± 0.01, 1.12 ± 0.01, 1.05 ± 0.01	1.42 ± 0.02, 1.20 ± 0.01, 1.22 ± 0.69
0.99	All	2.00 ± 0.01, 1.39 ± 0.01, 1.06 ± 0.01	2.70 ± 0.03, 1.29 ± 0.06, 1.15 ± 0.01	4.28 ± 0.45, 1.97 ± 0.95, 1.80 ± 0.29
	DEEP2	1.28 ± 0.01, 1.21 ± 0.01, 1.06 ± 0.01	1.57 ± 0.06, 1.29 ± 0.01, 1.16 ± 0.04	2.95 ± 0.63, 1.36 ± 0.76, 1.52 ± 0.68
	DESI	1.09 ± 0.01, 1.08 ± 0.01, 1.05 ± 0.01	1.28 ± 0.04, 1.17 ± 0.06, 1.08 ± 0.06	1.27 ± 0.80, 1.70 ± 0.96, 1.85 ± 1.09

Fig. 10 shows that the HOD of fixed number density samples of ELGs is very similar to that of SFR or L[O II]-selected samples, independently of the extra selection in either stellar mass, SFR, or L[O II]. This is because the ELGs we are studying are a subsample of star-forming galaxies. It is interesting to note that brighter [O II] emitters, have a reduced number of central galaxies in massive haloes, compared to fixed number density ELGs with an extra cut in SFR or stellar mass.

As it is shown in Fig. 10, the HOD of samples selected by their SFR or L[O II] have a larger number of central galaxies with low masses and a lower number of satellites at larger masses, when compared with stellar-mass-selected samples. This difference gets larger for decreasing number densities. Such difference is reduced for the ELG samples with fixed number densities.

Fig. 10 also shows that fixed number density L[O II]-selected samples populate slightly less massive haloes than the SFR-selected samples. This difference is reduced for the fixed number density ELG samples, in particular for lower number densities.

Although Fig. 10 only shows the results for the fixed number density VVDS-DEEP samples, similar trends are found for the other fixed number density ELG selections explored here, except for the DESI one. In this case, applying the different cuts result in minimal variations.

Compared to the global population, the minimum halo masses needed to host an ELG selected with a fixed number density are closer among the the cuts using the three properties studied here, stellar mass, SFR, and L[O II] . Despite the similarities, the effective bias of these samples are different, as it is described in the next section, Section 4.1.2.

For the fixed number density stellar mass selections, the minimum halo mass needed to find a galaxy increases for smaller number densities. This trend is also reflected in the increase of the effective bias from fixed number density stellar-mass-selected samples, as it can be seen in Table 5. This is not as clear for fixed number density SFR and L[O II] selections.

4.1.2 Clustering of fixed number density samples

The real-space two-point correlation function for galaxies with a fixed number density of $10^{-2}h^3$ Mpc^{-3} at redshift $z = 0.99$, is shown in Fig. 11. The calculation of the two-point correlation function has been done following the description in Section 3.2.2.

Fig. 11 shows that at large scales, $r > 8h^{-1}$ Mpc, the two-point correlation function of fixed number dentisy SFR and L[O II]-selected galaxies remain close, independently of starting with the whole galaxy population or ELGs. In fact, at a given number density, the bias of SFR and L[O II] cut samples for all galaxies and ELGs at $z = 0.83$ and $z = 0.99$ remain within a 0.6 range (0.3 if only number densities above $10^{-4.2}h^3\,\mathrm{Mpc}^{-3}$ are considered). The bias

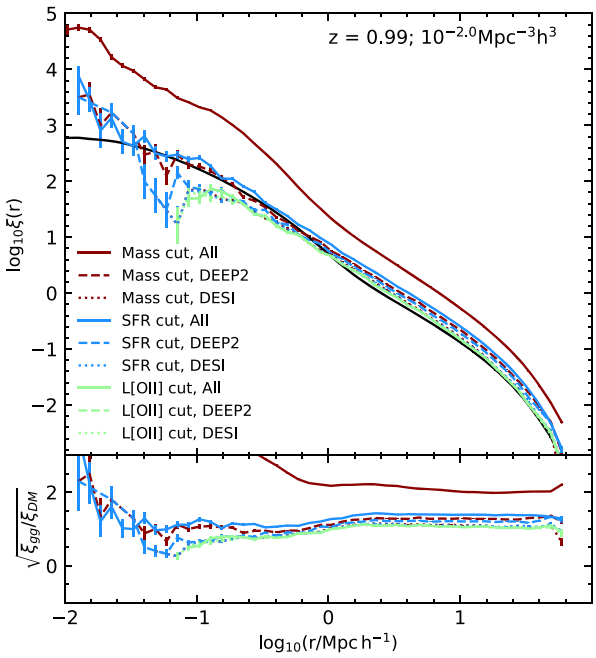

Figure 11. The real-space two-point correlation function for model galaxies with a number density of $10^{-2}h^3\mathrm{Mpc}^{-3}$ at redshift $z = 0.99$, selected using different criteria, as indicated in the legend. The black solid line shows the dark matter correlation function. *Bottom panel:* The real space ratio $\sqrt{\xi_{gg}/\xi_{DM}}$. Poisson error bars are shown in both panels.

of all studied samples can be seen in Table 5. The bias has been calculated in the range $8 \leq r(h^{-1}\,\mathrm{Mpc}) \leq 50$, as $\sqrt{\xi_{gg}/\xi_{DM}}$.

Both Fig. 11 and Table 5 show that galaxies selected with a single cut in stellar mass are more clustered than the rest of the samples (the solid red line versus the rest in Fig. 11). Although this is also true at large scales for fixed number density ELGs selected with an extra cut in stellar mass, in these cases the differences are much smaller (see the dashed and dotted red lines versus the blue ones in Fig. 11, for the case of DEEP2-selected galaxies). There is one exception to this: the least dense DESI sample, for which all selections are consistent at large scales.

Table 5 shows that, except for the DESI sample, the bias of mass-selected galaxies grow with lower number densities. Such a trend does not seem to exist for the other galaxy selections.

At large scales, the SFR and L[O II] cut samples trace closely the dark matter clustering, with biases, between 0.95 and 1.4 for samples with number densities above $10^{-4.2}h^3$ Mpc^{-3} (see Table 5). For these number densities, the L[O II] sample has bias slightly lower than the SFR one, being closer to 1. The clustering in the lowest studied number density bin becomes very noisy and despite the biases

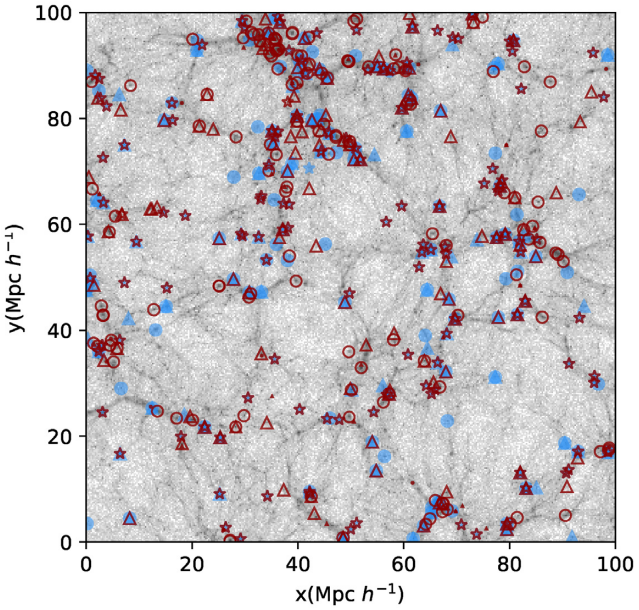

Figure 12. Distribution of galaxies with a number density of $10^{-2}h^3\,\mathrm{Mpc}^{-3}$, on top of the smooth underlying dark matter distribution (grey). This slice of $10h^{-1}$ Mpc thickness is taken from the MS-W7 simulation at $z = 0.99$. The red symbols show fixed number density galaxies selected with a single cut in stellar mass (the circles), SFR (the triangles), or L[O II] (stars), while the blue ones show those fixed number density DEEP2 galaxies selected with an extra cut on the mentioned properties. The area of the symbols is proportional to the $\log_{10}(L_{[\mathrm{OII}]})$.

reaching values close to 2, their corresponding error bars are close to 1.

The bias of galaxies selected with a single cut in either SFR or L[O II] is comparable to that of ELGs with a fixed number density.

As shown in Fig. 11, at small scales, the clustering of ELGs and galaxies selected with a single cut in SFR are different, except for the DESI-like sample. The clustering of this sample is consistent within the Poisson error bars for both mass and SFR-selected samples.

Fig. 11 shows that pairs of galaxies selected by their L[O II] are not found at the shortest separations found for the stellar mass or SFR samples. This is the case for all the studied selections. At $z = 0.99$ and $10^{-2}h^3\,\mathrm{Mpc}^{-3}$ (shown in Fig. 11), no pairs of L[O II]-selected galaxies are found with separations $r \lesssim 0.04h^{-1}$ Mpc. At $z = 0.83$ for the same number density, no L[O II]-selected galaxies are found with $r \lesssim 0.03h^{-1}$ Mpc. This values increase for decreasing number densities, a trend also seen for the other selections. The difference seen in Fig. 11 is striking. This difference is smaller for VVDS-DEEP at $z = 0.83$ and for lower number density samples. Nevertheless, this difference is worth exploring as its origin is unclear. In Fig. 9, different galaxy selections are compared in the SFR–stellar mass plane. From here, it is clear that very different ranges of stellar mass and SFR are covered by the different samples. These will cause differences in the clustering. We defer to an other study the intrahalo analysis of these samples, needed to better understand the differences in the clustering of galaxies within the same halo, the one-halo term.

4.2 Fixed number density samples in the cosmic web

Here, we study the fixed number density samples constructed in Section 4, to understand ELGs compared to mass- and SFR-selected samples within the cosmic web. Fig. 12 presents a 100 × 100 × $10h^{-3}\,\mathrm{Mpc}^3$ slice of the whole simulation box at redshift $z = 0.99$, highlighting in grey the cosmic web of the dark matter, together with the location of galaxies with a fixed number density, $10^{-2}h^3\mathrm{Mpc}^{-3}$, selected with single cuts on their stellar mass, SFR or L[O II] and ELGs with and extra cut on these properties, as described previously. Fig. 12 shows that, at least qualitatively, even when the number density is fixed, star-forming galaxies tend to trace less dense environments than mass-selected samples. In Fig. 12, it is unclear if there are significant differences between the large-scale environment traced by ELGs and galaxies selected by their SFR to have the same number density. In this section, we attempt to quantify the large-scale environment of ELGs and galaxies selected by their stellar mass and SFR.

4.2.1 *Large-scale environment distribution*

Following the methods described in Section 2.2, we classify the large-scale environment into voids, sheets, filaments, and knots using a velocity–shear–tensor algorithm, VWEB , with a 0.1 threshold, for the samples of galaxies with fixed number density constructed in Section 4. Fig. 13 compares selections in stellar mass, SFR and L[O II] with the same number density. For most of the galaxy selections considered, about half of these galaxies populate filaments. This is not the case for the mass selected sample with number densities below $10^{-2}h^3\,\mathrm{Mpc}^{-3}$ and for DEEP2 and VVDS-DEEP galaxies with an extra cut in stellar mass to achieve the lowest number density studied here, $10^{-4.2}h^3\,\mathrm{Mpc}^{-3}$.

As expected, samples based on a single stellar mass cuts have a higher presence in knots than the rest of the selections, which are star-forming galaxies. For all samples, the presence in knots increases for lower number densities.

L[O II]-selected samples trace the same large-scale environment structures, independently of being selected with just a single cut in L[O II] or not. This stresses that, at least for the studied number densities, the particular magnitude and colour cuts applied to select ELGs are secondary to the L[O II] limits.

L[O II]-selected samples are, in general, more present in sheets, ∼30 per cent, than galaxies selected with a single cut in SFR. For number densities below $10^{-2}h^3\,\mathrm{Mpc}^{-3}$, L[O II] galaxies are about 5 per cent more present in sheets, and less present in knots, than ELGs selected in other ways. This is accordant with the difference found for the clustering of their the one-halo term, reported in Section 4.1.2.

All the studied ELG selections are distributed in the cosmic web close to that of samples with the same number density based on a single SFR cut for number densities above $10^{-4.2}h^3\,\mathrm{Mpc}^{-3}$, with differences below a 0.11 ratio. ELGs selected with a number density of $10^{-2}h^3\,\mathrm{Mpc}^{-3}$ are about 5 per cent more present in sheets than the sample selected only by a cut in SFR.

The ELG sample with an extra cut in SFR to fix the number density, closely follow the distribution of the SFR sample, with differences in fractions up to 0.07. The differences between the ELG sample with an extra cut in stellar mass and the SFR one increase with decreasing number densities. At the lowest studied number density, the ELGs with an extra stellar mass cut have a much larger presence in knots than the SFR sample, with difference in fractions up to 0.48.

Above 70 per cent of the model ELG samples with number densities $10^{-2}h^3\mathrm{Mpc}^{-3}$ and $10^{-3}h^3\mathrm{Mpc}^{-3}$ are found in either filaments or sheets and about half of them are indeed in filaments. For the samples with a number density of $10^{-4.2}h^3\mathrm{Mpc}^{-3}$, this is only true for ELG samples with an extra cut in SFR or L[O II]; for stellar-mass-selected ELG samples, the percentage drops for all the ELGs, except for the eBOSS-SGC.

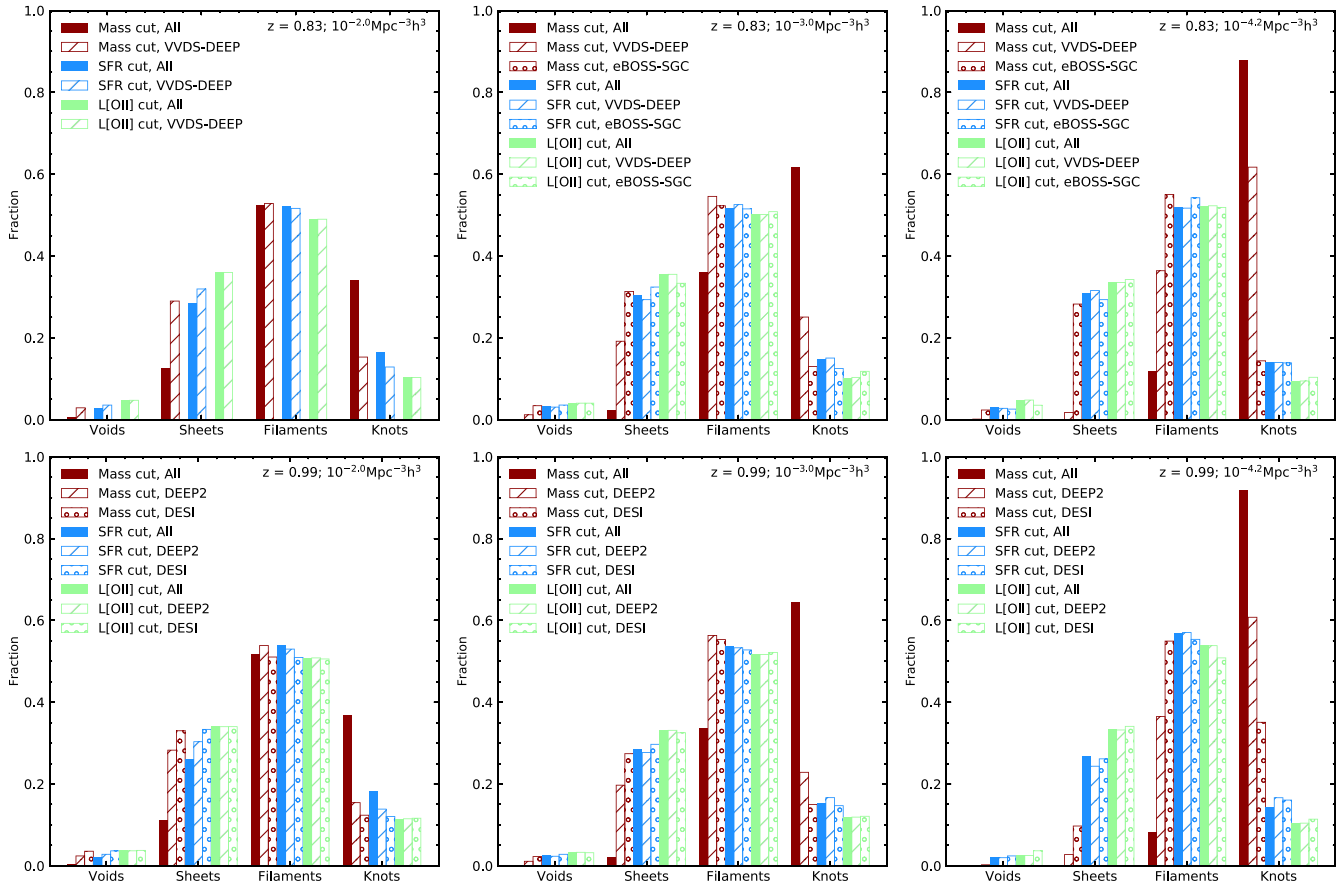

Figure 13. Histograms with the fraction of fixed number density samples of galaxies in voids, sheets, filaments, and knots. The large-scale environment has been classified using VWEB (see Section 2.2 for details). The top row show galaxies at $z = 0.83$ and the bottom row at $z = 0.99$. Selections with number densities of $10^{-2}h^3\,\mathrm{Mpc}^{-3}$ are shown in the left column, with $10^{-3}h^3\,\mathrm{Mpc}^{-3}$ in the middle column and with $10^{-4.2}h^3\,\mathrm{Mpc}^{-3}$ in the right column. Each panel shows galaxies selected with a single stellar mass, SFR, or L[O II] cut and those adding these cuts to the ELG selections, as indicated in the legend. Note that there are no eBOSS-SGC galaxies with number densities of $10^{-2}h^3\,\mathrm{Mpc}^{-3}$ at $z = 0.83$, as it can be seen in Fig. 8.

The environmental split does agree with the differences in the clustering amplitudes reported in Section 4.1.2. Fig. 11 shows that when more galaxies are found in knots the one-halo term of the two-point correlation function is much higher than that of the dark matter.

We have done a similar analysis but classifying the large-scale environment with a tidal–tensor algorithm, PWEB, with a $0.005\,s^{-2}$ threshold. The results with PWEB are quantitatively similar to those described above and can be seen in the Appendix C.

ELGs and L[O II]-selected galaxies tend to occupy either filaments or sheets. ELGs and L[O II]-selected galaxies roughly populate the same large-scale environment as galaxies selected based on their SFR, for number densities above or equal to $10^{-3}h^3\mathrm{Mpc}^{-3}$. Below this number density, the differences can be large for stellar-mass-selected ELGs, in particular in knots and voids.

4.2.2 Comparison of global properties

For a given galaxy sample, the median stellar mass is comparable for galaxies in knots, filaments, sheets, and voids. As expected, the maximum stellar mass of galaxies decreases towards less dense environments, i.e. it decreases from knots to voids. A similar trend is seen for the host halo mass of the galaxies, shown in the left-hand panel in Fig. 14 for galaxies selected at $z = 0.99$ with $10^{-3}h^3\mathrm{Mpc}^{-3}$. This trend does affect the distribution of host halo masses, such that median halo masses can decrease from knots to voids.

The median SFR is comparable for a given sample of galaxies in knots, filaments, and sheets. In voids, stellar-mass-selected galaxies have minimum SFR above those for knots, filaments, and Sheet. For example, galaxies with $10^{-3}h^3\,\mathrm{Mpc}^{-3}$ at $z = 0.99$ in voids have SFR $> 10^{9}h^{-1}\,\mathrm{M}_{\odot}\,\mathrm{Gyr}^{-1}$, while in the other large-scale structures galaxies with SFR $< 10^{7}h^{-1}\,\mathrm{M}_{\odot}\,\mathrm{Gyr}^{-1}$ can be found. This suggests that galaxies in voids at $z \sim 1$ are less affected by the quenching of the star formation than in other large-scale environment. A similar difference between voids and the other large-scale structures is found for the distribution of mass-selected samples as a function of sSFR, as shown in the middle panel of shown in Fig. 14 for galaxies selected with $10^{-3}h^3\,\mathrm{Mpc}^{-3}$ at $z = 0.99$. However, the distribution as a function of sSFR of galaxies selected with a single stellar mass cut has a larger variation with the large-scale structure. In Fig. 14, the differences between filaments, sheets, and voids are clear for galaxies selected with a single cut in stellar mass. This trend is in agreement with star formation being quenched more effectively in the densest large-scale structures for a mass-selected sample. This is not as clear for star-forming galaxies, for which a minimum SFR or L[O II] has already been imposed.

The right-hand panel in Fig. 14 shows as a function of L[O II] the distribution of galaxies with $10^{-3}h^3\,\mathrm{Mpc}^{-3}$ at $z = 0.99$. The distributions are comparable for galaxies in different large-scale structures.

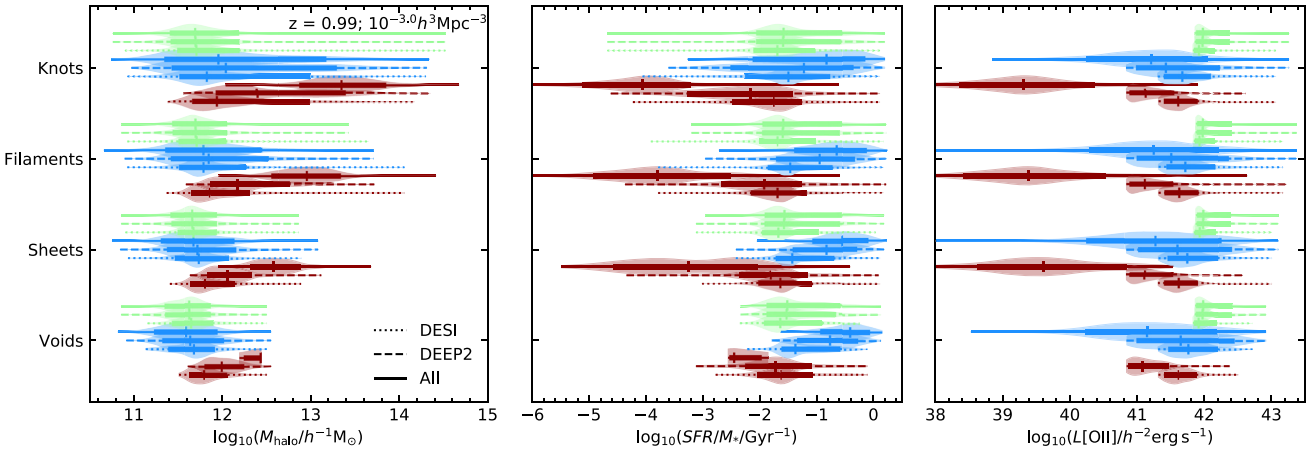

Figure 14. Distribution of the host halo mass, left-hand panel, specific SFR, middle panel, and L[O II], right-hand panel, of galaxies with a fixed number density of $10^{-3}h^3Mpc^{-3}$ at $z = 0.99$, in the knots, filaments, sheets, and voids classifications of the large-scale environment done with the VWEB algorithm. The distributions of L[O II]-selected samples is shown in green, SFR-selected samples in blue, and those for stellar-mass-selected samples in red. The horizontal lines indicate the extent of the variables for each galaxy selection. The solid lines correspond to selections made with a single cut in either stellar mass or SFR, while those on top of the ELG selections summarized in Table 2 are shown by the dashed or dotted lines as indicated in the legend. For each selection, the vertical line indicates its median, the thick horizontal line the range of the 10–90 per cent of the distribution and the violin type areas show the data distribution using a Gaussian kernel density estimation. Note that galaxies selected with a single mass cut have no [O II] emission in voids.

The trends discussed above are found for the classifications of the large-scale environment done with both the VWEB and PWEB algorithms.

The percentages of satellites in different large-scale environment with $10^{-3}h^3$ Mpc^{-3} are summarized in Table 6. The percentage of satellite galaxies decreases with number density, as rarer objects are more likely to be central galaxies. For all the studied number densities, less than 10 per cent of model galaxies in either voids, sheets, or filaments are satellites. In knots, this percentage increases and it can go up to 30 per cent for stellar mass and SFR-selected samples with a fixed number density of $10^{-2}h^3$ Mpc^{-3}. The percentage of satellites within the L[O II]-selected sample remains below 15 per cent in all studied cases. Groups and clusters of galaxies tend to occur in knots and thus, the fraction of satellite galaxies is expected to be larger there (e.g Guo, Tempel & Libeskind 2015). We find here that this trend is maintained also for star-forming galaxies, although the variation is slightly smaller, as can be seen in Table 6. The percentage of satellite galaxies in different large-scale environment are comparable between the VWEB and PWEB classifications.

Table 6. Percentage of satellite galaxies for the fixed number density samples with $10^{-3}h^3$ Mpc^{-3} (see Section 4), at $z = 0.83$ and $z = 0.99$. The percentage of satellites is tabulated for the stellar mass, SFR, and L[O II]-selected samples separated by a comma, for each large-scale environment classified using the VWEB algorithm (see Section 4.2).

$z = 0.83$	All	VVDS-DEEP	eBOSS-SGC
Voids	0.0, 1.8, 1.6	0.7, 1.8, 1.6	1.6, 1.3, 0.8
Sheets	0.3, 1.4, 1.7	1.3, 1.3, 1.7	2.6, 2.2, 2.4
Filaments	2.5, 2.3, 3.4	3.7, 2.2, 3.4	5.2, 4.5, 4.4
Knots	8.7, 4.3, 6.1	8.8, 4.5, 6.0	14.3, 12.4, 10.2
$z = 0.99$	All	DEEP2	DESI
Voids	0.0, 1.6, 1.7	0.7, 0.7, 1.7	0.7, 0.3, 1.5
Sheets	0.6, 1.4, 1.6	0.9, 1.5, 1.6	0.9, 1.3, 1.7
Filaments	2.1, 2.6, 2.8	2.6, 2.2, 2.8	2.7, 2.4, 3.0
Knots	7.9, 5.4, 7.3	6.2, 4.4, 7.3	5.4, 5.1, 7.5

5 SUMMARY AND CONCLUSIONS

Star-forming ELGs are being targeted by current and future cosmological redshift surveys. Here, we have studied how they populate the cosmic web structure according to an SAM of galaxy formation and evolution. In the future, we expect observational studies to characterize how ELGs trace the cosmic web. Here, we have also contrasted ELGs with samples selected in simpler, and perhaps more generic, ways, by ranking galaxies in properties such as stellar mass or SFR and applying a cut.

We have used a new flavour of the SAM GALFORM run on the MS-W7 dark-matter-only simulation, with a WMAP7 cosmology and a simulation box of side $500h^{-1}$Mpc. This new version improves on the model presented in Gonzalez-Perez et al. (2018) by (i) stripping the gas in satellites more slowly, such that the observed passive fraction at $z = 0$ is better matched (see Fig. 2 in Section 2.1) and (ii) including the updated treatment of the evolution of SMBHs introduced in Griffin et al. (2019). The last point is relevant for this work as this improvement results in a different evolution of the AGN feedback. The model has been calibrated against local observations.

Model ELGs (driven by star formation rather than nuclear activity) are selected by imposing cuts on apparent magnitude and [O II] flux to mimic five observational surveys: DEEP2, VVDS-Deep, VVDS-Wide, eBOSS-SGC, and DESI (see Table 2). Further colour cuts are imposed in the latter two cases to mimic the spectroscopic selection that avoids targeting objects with colours that can be confused with stars.

The large-scale bias of model ELGs is close to unity (see Fig. 7 in Section 3.2.2). These model galaxies are naturally affected by assembly bias, as the SAM of galaxy formation and evolution includes the effect of different halo assembly histories. Thus, the bias measured from galaxy catalogues constructed with a HOD model without considering assembly bias, might be slightly larger than the values reported here.

The large-scale environment at $z = 0.83,\ 0.99$ has been classified into voids, sheets, filaments, and knots using (i) a velocity–shear–tensor algorithm, VWEB, with a threshold of 0.1 and (ii) a tidal–tensor algorithm, PWEB, with a threshold of 0.005 s^{-2}. Similar conclusions are reached with both algorithms.

Half of the model ELGs live in filaments and a third in sheets (see Section 3.2). Model ELGs in knots have the largest percentage of satellite galaxies and a tail of massive galaxies that sets them apart when comparing the galaxy stellar mass function in each environment. We find that the shape of the mean HOD of model ELGs varies widely with large-scale environment (see Fig. 6), partly due to the different presence of satellite ELGs in different cosmic web structures. The mean HOD of voids and sheets, where almost all galaxies are centrals, has a shape close to an asymmetric Gaussian. The mean HOD of central galaxies in filaments and knots has a plateau. The presence of satellite galaxies is most important among ELGs in knots, for which the mean satellite HOD follows a typical power law.

We have explored the cross-correlation between the whole ELG sample and those living in voids, sheets, filaments, or knots (see Fig. 7). We find that, for all the studied ELG samples, the clustering of ELGs in knots is boosted at $1 \leq r(h^{-1}\,\mathrm{Mpc}) \leq 10$, while ELGs in voids are less clustered on large scales, $r > 1h^{-1}\,\mathrm{Mpc}$.

To put in context the results obtained for model ELGs, we have defined samples with three fixed number densities, 10^{-2}, 10^{-3}, $10^{-4.2}h^3\,\mathrm{Mpc}^{-3}$. These samples have been selected by imposing an extra cut in either stellar mass, SFR, or L[O II] for ELGs or by imposing a single cut in one of these three properties to the whole sample of model galaxies, as shown in Fig. 8. The median L[O II] for galaxies selected only by their stellar mass is below $10^{39.5}h^{-2}\mathrm{erg\,s}^{-1}$, while all the other selections are much brighter, with median L[O II] above $10^{40.5}h^{-2}\mathrm{erg\,s}^{-1}$.

The mean HODs of model ELGs with fixed number densities have shapes close to those of star-forming samples, selected either based on a SFR or L[O II] cut (see Fig. 10). The studied ELGs are indeed a subsample of the star-forming population.

For a fixed number density, we find that, in general, star-forming galaxies are less clustered than stellar mass selected ones (see Fig. 11). Fixed number density ELG, SFR, and L[O II]-selected samples have very similar large-scale bias. However, their clustering differs below separations of $1h^{-1}$Mpc. For instance, no pairs of L[O II]-selected samples are found at the smallest separations considered. This might have implications for the expectations of redshift–space distortions derived assuming that ELGs are equivalent to galaxies selected by a single cut in SFR (e.g. Orsi & Angulo 2018; Jiménez et al. 2019).

As expected, fixed number density samples selected with a single stellar mass cut have a higher presence in knots than either ELGs or galaxies selected by their SFR or L[O II].

For a fixed number density, the distribution of star-forming ELGs in the cosmic web follows closely that of samples selected with a single cut in eitheir SFR or L[O II] (see Fig. 13). The differences are more significant for low number density samples, at least with respect to SFR selected samples.

Over 70 per cent of the model ELG samples with number densities $10^{-2}h^3\,\mathrm{Mpc}^{-3}$ and $10^{-3}h^3\,\mathrm{Mpc}^{-3}$ are found in either filaments or sheets. About half of them are in filaments. For samples with lower number densities, this percentage drops, except for the eBOSS-SGC model sample.

The maximum stellar mass and host halo mass decreases from knots to voids for both star-forming and stellar-mass-selected samples with fixed number densities (see Fig. 14). The specific star formation of fixed number density model samples is largely independent of the large-scale environment for star-forming galaxies, but increases moving from knots to voids for galaxies selected with a single cut in stellar mass. For a fixed number density model sample, the L[O II] appears to be independent of the large-scale environment.

The agreement between the properties of the ELGs, SFR, and L[O II]-selected samples, at least for number densities above $10^{-4.2}h^3\,\mathrm{Mpc}^{-3}$, shows the robustness of our results. For large scales, one could use the dispersion in the two-point correlation function among these 'star-forming' samples, as a reasonable estimate of the systematic error when producing mock catalogues. For small scales, variations are found among star-forming galaxies selected in different ways.

ACKNOWLEDGEMENTS

The authors would like to thank the help provided by Nuala McCullagh, Lee Stothert, and Alex Smith to run CUTE on hdf5 files. VGP acknowledges past support from the University of Portsmouth through the Dennis Sciama Fellowship award. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant 769130). WC was supported by the *Ministerio de Economía y Competitividad* and the *Fondo Europeo de Desarrollo Regional* (MINECO/FEDER, UE) in Spain through grant AYA2015-63810-P. WC further acknowledges the support from the European Research Council under grant 670193. SC acknowledges the support from the Juan de la Cierva Formación Fellowship (FJCI-2017-33816). AG acknowledges support from the STFC studentship ST/N50404X/1. CGL and CMB acknowledge support from the STFC grant ST/P000541/1. AK is supported by MINECO/FEDER (Spain) under research grants AYA2015-63819-P and PGC2018-094975-C2. He further acknowledges support from the Spanish Red Consolider MultiDark FPA2017-90566-REDC and thanks Tocotronic for the red album. PN acknowledges the support of the Royal Society through the award of a University Research Fellowship, and the European Research Council, through receipt of a Starting grant (DEGAS-259586). This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/P002293/1, ST/R002371/1, and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure. This work has benefitted from the publicly available programming language PYTHON (https://www.python.org/) and the package MATPLOTLIB (https://matplotlib.org/).

6 DATA AVAILABILITY

The programs and a small fraction of the data used to generate the plots presented in this paper can be found in https://github.com/viogp/plots4papers/tree/master/elg_cw_plots. Other sets of data can be shared upon request.

REFERENCES

Ahumada R. et al., 2019, ApJS, 249, 3
Alam S., Peacock J. A., Kraljic K., Ross A. J., Comparat J., 2020, MNRAS, 497, 581
Alonso D., 2012, preprint (arXiv:1210.1833)
Alpaslan M. et al., 2016, MNRAS, 457, 2287
Anders P., Fritze-v. Alvensleben U., 2003, A&A, 401, 1063
Avila S. et al., 2020, preprint (arXiv:2007.09012)
Bauer A. E. et al., 2013, MNRAS, 434, 209
Baugh C. M., 2006, Rep. Prog. Phys., 69, 3101
Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
Baugh C. M. et al., 2018, MNRAS, 483, 4922
Benson A. J., 2010, Phys. Rep., 495, 33
Benson A. J., Bower R., 2010, MNRAS, 405, 1573
Bond J. R., Kofman L., Pogosyan D., 1996, Nature, 380, 603

Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Carlesi E., Knebe A., Lewis G. F., Wales S., Yepes G., 2014, MNRAS, 439, 2943
Cautun M., van de Weygaert R., Jones B. J. T., Frenk C. S., 2014, MNRAS, 441, 2923
Chen Y.-C. et al., 2017, MNRAS, 466, 1880
Cochrane R. K., Best P. N., 2018, MNRAS, 480, 864
Cochrane R. K., Best P. N., Sobral D., Smail I., Wake D. A., Stott J. P., Geach J. E., 2017, MNRAS, 469, 2913
Cochrane R. K., Best P. N., Sobral D., Smail I., Geach J. E., Stott J. P., Wake D. A., 2018, MNRAS, 475, 3730
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Colless M. et al., 2001, MNRAS, 328, 1039
Comparat J. et al., 2013a, MNRAS, 428, 1498
Comparat J. et al., 2013b, MNRAS, 428, 1498
Comparat J. et al., 2013c, MNRAS, 433, 1146
Comparat J. et al., 2015, A&A, 575, A40
Comparat J. et al., 2016a, MNRAS, 461, 1076
Comparat J. et al., 2016b, A&A, 592, A121
Comparat J. et al., 2017, preprint (arXiv e-prints)
Contreras S., Baugh C. M., Norberg P., Padilla N., 2013, MNRAS, 432, 2717
Contreras S., Zehavi I., Padilla N., Baugh C. M., Jiménez E., Lacerna I., 2019, MNRAS, 484, 1133
Cui W., Borgani S., Dolag K., Murante G., Tornatore L., 2012, MNRAS, 423, 2279
Cui W., Borgani S., Murante G., 2014, MNRAS, 441, 1769
Cui W., Knebe A., Yepes G., Yang X., Borgani S., Kang X., Power C., Staveley-Smith L., 2018, MNRAS, 473, 68
Cui W. et al., 2019, MNRAS, 485, 2367
Darvish B., Sobral D., Mobasher B., Scoville N. Z., Best P., Sales L. V., Smail I., 2014, ApJ, 796, 51
Darvish B., Mobasher B., Sobral D., Hemmati S., Nayyeri H., Shivaei I., 2015, ApJ, 814, 84
Dawson K. S. et al., 2016, AJ, 151, 44
de Jong R. S. et al., 2014, The Messenger, 175, 3
Delubac T. et al., 2017, MNRAS, 465, 1831
DESI Collaboration, 2016, preprint (arXiv:1611.00036)
Drinkwater M. J. et al., 2010, MNRAS, 401, 1429
Drinkwater M. J. et al., 2018, MNRAS, 474, 4151
Driver S. P. et al., 2012, MNRAS, 427, 3244
Fanidakis N., Baugh C. M., Benson A. J., Bower R. G., Cole S., Done C., Frenk C. S., 2011, MNRAS, 410, 53
Favole G. et al., 2016, MNRAS, 461, 3421
Font A. S. et al., 2008, MNRAS, 389, 1619
Franx M., van Dokkum P. G., Förster Schreiber N. M., Wuyts S., Labbé I., Toft S., 2008, ApJ, 688, 770
Furlong M. et al., 2015, MNRAS, 450, 4486
Geach J. E., Sobral D., Hickox R. C., Wake D. A., Smail I., Best P. N., Baugh C. M., Stott J. P., 2012, MNRAS, 426, 679
Geller M. J., Huchra J. P., 1989, Science, 246, 897
Gilbank D. G., Baldry I. K., Balogh M. L., Glazebrook K., Bower R. G., 2010, MNRAS, 405, 2594
Goh T. et al., 2019, MNRAS, 483, 2101
Gonzalez-Perez V., Baugh C. M., Lacey C. G., Kim J.-W., 2011, MNRAS, 417, 517
Gonzalez-Perez V., Lacey C. G., Baugh C. M., Frenk C. S., Wilkins S. M., 2013, MNRAS, 429, 1609
Gonzalez-Perez V., Lacey C. G., Baugh C. M., Lagos C. D. P., Helly J., Campbell D. J. R., Mitchell P. D., 2014, MNRAS, 439, 264
Gonzalez-Perez V. et al., 2018, MNRAS, 474, 4024 (GP18)
Gott J. Richard I., Jurić M., Schlegel D., Hoyle F., Vogeley M., Tegmark M., Bahcall N., Brinkmann J., 2005, ApJ, 624, 463
Griffin A. J., Lacey C. G., Gonzalez-Perez V., Lagos C. d. P., Baugh C. M., Fanidakis N., 2019, MNRAS, 487, 198
Guo H. et al., 2018a, ApJ, 871, 147
Guo H. et al., 2019, preprint (arXiv e-prints)
Guo Q., White S., Angulo R. E., Henriques B., Lemson G., Boylan-Kolchin M., Thomas P., Short C., 2013, MNRAS, 428, 1351
Guo Q., Tempel E., Libeskind N. I., 2015, ApJ, 800, 112
Hahn O., Porciani C., Carollo C. M., Dekel A., 2007, MNRAS, 375, 489
Hayashi M. et al., 2020, PASJ, tmp 232
Hill G. J. et al., 2008, in Kodama T., Yamada T., Aoki K., eds, ASP Conf. Ser. Vol. 399, Panoramic Views of Galaxy Formation and Evolution. Astron. Soc. Pac., San Fransisco, p. 115
Hoffman Y., Metuki O., Yepes G., Gottlöber S., Forero-Romero J. E., Libeskind N. I., Knebe A., 2012, MNRAS, 425, 2049
Hounsell R. et al., 2018, ApJ, 867, 23
Jiang L., Helly J. C., Cole S., Frenk C. S., 2014, MNRAS, 440, 2115
Jiménez E., Contreras S., Padilla N., Zehavi I., Baugh C. M., Gonzalez-Perez V., 2019, MNRAS, 490, 3532
Khostovan A. A. et al., 2018, MNRAS, 478, 2999
King A. R., Pringle J. E., Hofmann J. A., 2008, MNRAS, 385, 1621
Klypin A. A., Shandarin S. F., 1983, MNRAS, 204, 891
Komatsu E. et al., 2011, ApJS, 192, 18
Kraljic K. et al., 2018, MNRAS, 474, 547
Kraljic K. et al., 2019, MNRAS, 483, 3227
Lacey C. G. et al., 2016, MNRAS, 462, 3854
Lagos C. D. P., Davis T. A., Lacey C. G., Zwaan M. A., Baugh C. M., Gonzalez-Perez V., Padilla N. D., 2014, MNRAS, 443, 1002
Lagos C. D. P., Lacey C. G., Baugh C. M., Bower R. G., Benson A. J., 2011, MNRAS, 416, 1566
Laigle C. et al., 2015, MNRAS, 446, 2744
Laigle C. et al., 2018, MNRAS, 474, 5437
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Le Fèvre O. et al., 2013, A&A, 559, A14
Leauthaud A. et al., 2007, ApJS, 172, 219
Levi M. et al., 2013, preprint (arXiv:1308.0847)
Liao S., Gao L., 2019, MNRAS, 485, 464
Libeskind N. I., Hoffman Y., Knebe A., Steinmetz M., Gottlöber S., Metuki O., Yepes G., 2012, MNRAS, 421, L137
Libeskind N. I., Hoffman Y., Forero-Romero J., Gottlöber S., Knebe A., Steinmetz M., Klypin A., 2013, MNRAS, 428, 2489
Libeskind N. I. et al., 2018, MNRAS, 473, 1195
Luber N., van Gorkom J. H., Hess K. M., Pisano D. J., Fernández X., Momjian E., 2019, AJ, 157, 254
Malavasi N. et al., 2017, MNRAS, 465, 3817
McCarthy I. G., Frenk C. S., Font A. S., Lacey C. G., Bower R. G., Mitchell N. L., Balogh M. L., Theuns T., 2008, MNRAS, 383, 593
McConnell N. J., Ma C.-P., 2013, ApJ, 764, 184
Mitchell P. D. et al., 2018, MNRAS, 474, 492
Newman J. A. et al., 2013, ApJS, 208, 5
Norberg P. et al., 2002, MNRAS, 336, 907
Orsi A., Lacey C. G., Baugh C. M., Infante L., 2008, MNRAS, 391, 1589
Orsi Á. A., Angulo R. E., 2018, MNRAS, 475, 2530
Paulino-Afonso A., Sobral D., Darvish B., Ribeiro B., Smail I., Best P., Stroe A., Cairns J., 2019, A&A, 633, A70
Percival W. J. et al., 2019, preprint (arXiv:1903.03158)
Planck Collaboration XVI, 2014, A&A, 571, A16
Poudel A., Heinämäki P., Tempel E., Einasto M., Lietzen H., Nurmi P., 2017, A&A, 597, A86
Raichoor A. et al., 2017, MNRAS, 471, 3955
Sinha M., Garrison L. H., 2020, MNRAS, 491, 3022
Sobral D., Best P. N., Matsuda Y., Smail I., Geach J. E., Cirasuolo M., 2012, MNRAS, 420, 1926
Sobral D., Stroe A., Koyama Y., Darvish B., Calhau J., Afonso A., Kodama T., Nakata F., 2016, MNRAS, 458, 3443
Somerville R. S., Davé R., 2015, ARA&A, 53, 51
Stasińska G., 1990, A&AS, 83, 501
Takada M. et al., 2014, PASJ, 66, R1
Valentino F. et al., 2017, MNRAS, 472, 4878
Vulcani B. et al., 2019, MNRAS, 487, 2278
Wang Y. et al., 2018, Publ. Astron. Soc. Aust., 36, e015
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Zentner A. R., Hearin A. P., van den Bosch F. C., 2014, MNRAS, 443, 3044
Zheng Z. et al., 2005, ApJ, 633, 791

APPENDIX A: EBOSS-SGC AND DESI COLOUR CUTS

Fig. A1 presents the location of model galaxies with $\mathrm{Flux}_{[\mathrm{OII}]} > 8 \times 10^{-17}\mathrm{erg\,s^{-1}cm^{-2}}$ at redshifts $z = 0.62, 0.83, 1., 1.5$ in the $(g-r)_{\mathrm{DECam}}$ versus $(r-z)_{\mathrm{DECam}}$, colour–colour space. These distributions are compared to the location of stars (Leauthaud et al. 2007), the grey-filled symbols in Fig. A1 and to the regions delimited by the eBOSS-SGC (Raichoor et al. 2017) and DESI colour cuts DESI Collaboration (2016). These colour cuts are summarized in Table 2.

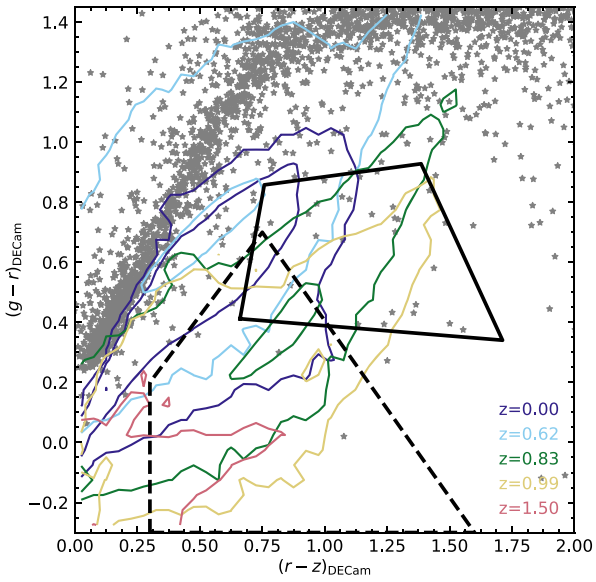

Figure A1. DECam $(g-r)$ versus $(r-z)$ parameter space with the isodensity lines at $\log_{10}(\Phi/Mpc^{-3}h^3\mathrm{dlog_{10}L}) = -4.5, -1.5, 1$ for model galaxies, with $\mathrm{Flux}_{[\mathrm{O\,II}]} > 8 \times 10^{-17}\mathrm{erg\,s^{-1}cm^{-2}}$ at the redshifts indicated in the legend. The polygon with the solid lines shows the eBOSS-SGC colour selection and the one with dashed lines that for DESI, as summarized in Table 2. Note that the flux limit difference between the eBOSS-SGC and DESI selections is about a 20 per cent and thus, the distribution of model galaxies in this plot is very similar for both cuts, $\mathrm{Flux}_{[\mathrm{O\,II}]} > 8 \times 10^{-17}\mathrm{erg\,s^{-1}cm^{-2}}$ and $\mathrm{Flux}_{[\mathrm{O\,II}]} > 10^{-16}\mathrm{erg\,s^{-1}cm^{-2}}$. The location of stars are shown by the grey symbols.

The colours of model galaxies are roughly consistent with the regions defined for eBOSS-SGC (Comparat et al. 2016b) and DESI (DESI Collaboration 2016) to select ELGs in the range $0.6 < z < 1.7$. Further details on the colour cuts can be found in Gonzalez-Perez et al. (2018).

APPENDIX B: RESOLUTION AND THRESHOLD CHECKING

In Fig. B1, we compare the results with two different mesh numbers 256^3 (corresponding to a cell size of $\sim 2\ h^{-1}$Mpc) and 512^3 ($\sim 1\ h^{-1}$Mpc cell size). Both meshes adopt the referenced thresholds: $\lambda_{\mathrm{th}}^V = 0.1$ for VWEB and $\lambda_{\mathrm{th}}^P = 0.005 s^{-2}$ for PWEB. The smoothing length in all cases is set to $5h^{-1}$Mpc (see Section 4.2). It is clear that more details are revealed with the finer meshes. However, we do not go beyond the 512^3 number of meshes, as finer mesh cells will have less particles which will provide noisier fields. We confirm here that with these two thresholds we have very similar volume fractions (see Tabel B1 for details) between VWEB and PWEB classified large-scale structures. Similar fractions are also found assuming $\lambda_{\mathrm{th}}^P = 0.00346 s^{-2}$. Furthermore, we can see that these classified large-scale structures with both methods match the density fields shown in the right-hand side panels well.

We further investigated the effects of varying the two thresholds within two times of the reference values. They are either too large – $\lambda_{\mathrm{th}}^V = 0.05$ ($\lambda_{\mathrm{th}}^P = 0.0025 s^{-2}$) – with more knots regions, or too small – $\lambda_{\mathrm{th}}^V = 0.2$ ($\lambda_{\mathrm{th}}^P = 0.1 s^{-2}$) – with more space is occupied by Void. However, it is interesting to see that our main conclusions are basically unchanged.

Table B1. The fraction in the simulation volume of each large-scale structures with different mesh numbers. Similar fractions were found in Cui et al. (2019).

Mesh number	Method	Knot	Filament	Sheet	Void
256^3	VWEB	0.026	0.209	0.469	0.296
256^3	PWEB	0.027	0.252	0.490	0.227
512^3	VWEB	0.029	0.227	0.474	0.270
512^3	PWEB	0.021	0.242	0.510	0.226

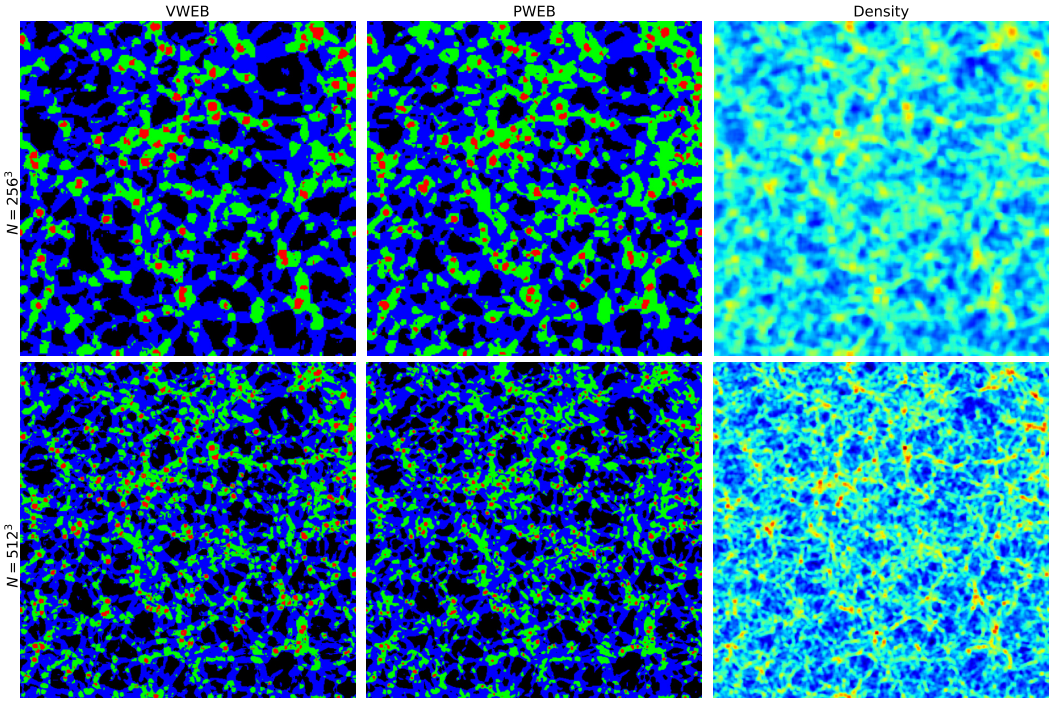

Figure B1. The projected structures for a slice of the simulation at $z = 1$ with 256^3 meshes (upper panels, $\sim 2\ h^{-1}$Mpc thickness) and 512^3 meshes (lower panels, $\sim 1\ h^{-1}$Mpc thickness). The large-scale structures: knots, filaments, sheets, and voids regions are shown as the red, green, blue, and black colours, respectively. From left-hand panel to right-hand panel, we show the results from VWEB and PWEB as well as the density fields. The thresholds applied are $\lambda_{\rm th} = 0.1$ for VWEB and $\lambda_{\rm th} = 0.005\ {\rm s}^{-2}$ for PWEB meshes.

APPENDIX C: LARGE-SCALE ENVIRONMENT WITH PWEB

Fig. C1 shows the fraction of galaxies in voids, sheets, filaments, and knots when the large-scale environment is classified using the PWEB algorithm described in Section 2.2, with a $0.005\ s^{-2}$ threshold. Very similar fractions are found when imposing 0.00346 and $0.01\ s^{-2}$ thresholds. This figure is qualitatively equivalent to Fig. 13, in terms of global trends. However, quantitatively there are differences that become more pronounced for the samples with the lowest number density, in particular for the mass selected ones.

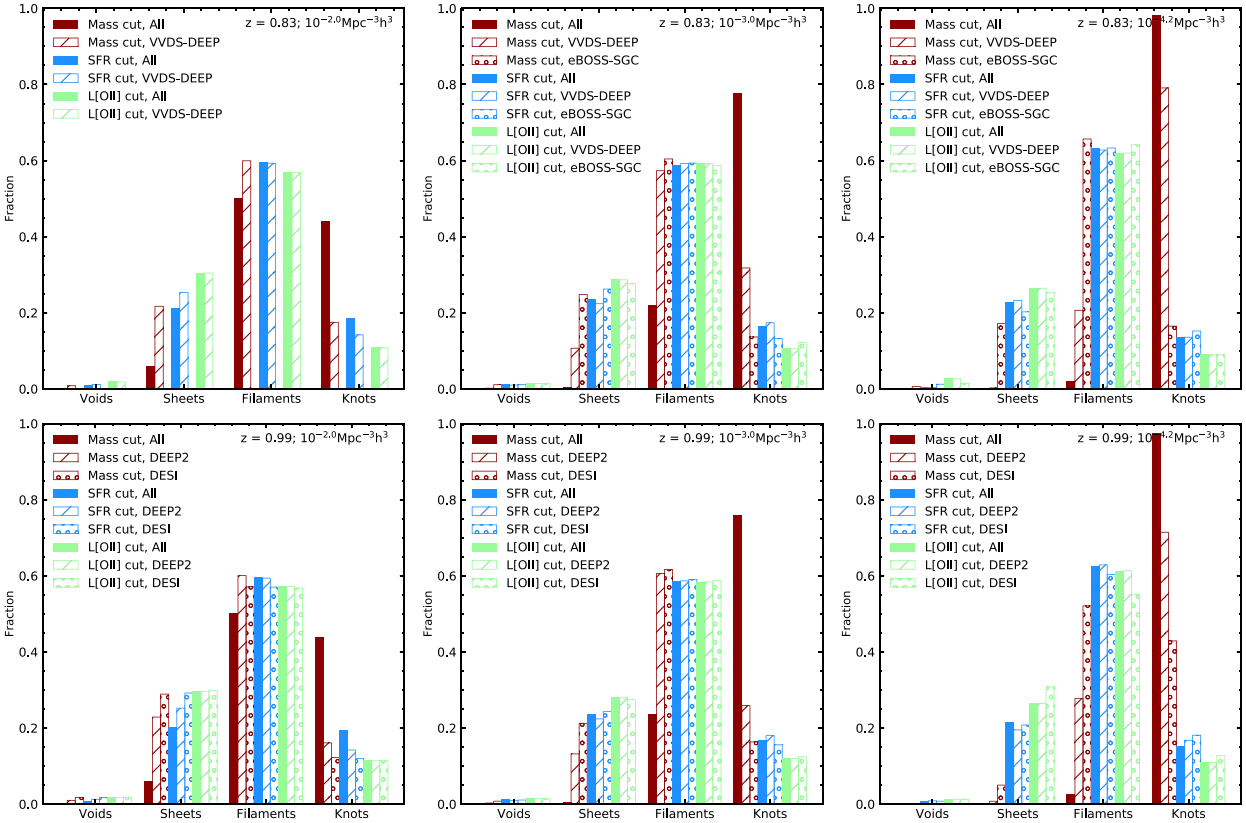

Figure C1. Histograms with the fraction of galaxies in voids, sheets, filaments, and knots, similar to Fig. 13. In this case, the large-scale environment has been classified using PWEB (see Section 2.2).

[1]*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK*
[2]*Energy Lancaster, Lancaster University, Lancaster LA14YB, UK*
[3]*Institute of Cosmology & Gravitation, University of Portsmouth, Dennis Sciama Building, Portsmouth PO1 3FX, UK*
[4]*Departamento de Física Teórica, Módulo 15, Facultad de Ciencias, Universidad Autónoma de Madrid, E-28049 Madrid, Spain*
[5]*Institute for Astronomy, University of Edinburgh, Royal Observatory, EH9 3HJ Edinburgh, UK*
[6]*Donostia International Physics Center (DIPC), Manuel Lardizabal Pasealekua 4, E-20018 Donostia, Basque Country, Spain*
[7]*Department of Physics, Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK*
[8]*Max-Planck Institut fur extraterrestrische Physik, Postfach 1312, D-85741 Garching bei Munchen, Germany*
[9]*Centro de Investigación Avanzada en Física Fundamental, Facultad de Ciencias, Universidad Autónoma de Madrid, E-28049 Madrid, Spain*
[10]*International Center for Radio Astronomy Research, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia*
[11]*Center for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*

This paper has been typeset from a TEX/LATEX file prepared by the author.